

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2004, 13:10:38 ; Search time 25.75 Seconds

(without alignments)
139.402 Million cell updates/sec

Title: SEQ-A

Perfect score: 98

Sequence: 1 HGHQQRGLGHGKXK1 17

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 809742 seqs, 21153259 residues

Total number of hits satisfying chosen parameters: 809742

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AB.*

1: /cgn2_6/ptodata/2/pubaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	96	98.0	415	15	US-10-162-335-76
2	96	98.0	615	15	US-10-162-335-72
3	96	98.0	644	15	US-10-162-335-74
4	96	98.0	644	15	US-10-162-335-84
5	60	61.2	574	14	US-10-156-761-1406
6	57	58.2	385	14	US-10-270-333-42
7	55	56.1	376	15	US-10-176-306-29
8	55	56.1	376	15	US-10-264-237-1632
9	55	56.1	278	14	US-10-176-306-59
10	54	55.1	503	9	US-09-738-897-5
11	54	55.1	507	9	US-09-738-897-4
12	54	55.1	507	10	US-09-849-138-41
13	54	55.1	507	14	US-10-176-306-49
14	53	54.1	469	15	US-10-360-849A-36
15	53	54.1	831	9	US-09-789-561-86

16	53	54.1	831	11	US-09-833-245-2153
17	51	52.0	91	13	US-10-090-035-22
18	51	52.0	92	13	US-10-090-035-20
19	51	52.0	92	13	US-10-090-035-24
20	49	50.0	69	14	US-10-156-761-8065
21	49	50.0	377	16	US-10-380-727-6
22	49	50.0	856	10	US-09-949-029-66
23	48	49.0	444	15	US-10-360-849A-30
24	47	48.0	123	14	US-10-106-698-6826
25	47	48.0	181	14	US-10-180-375-170
26	47	48.0	191	15	US-10-374-780A-797
27	46.5	47.4	476	15	US-10-360-849A-33
28	46	46.9	130	16	US-10-389-566-2423
29	46	46.9	133	16	US-10-389-566-2424
30	46	46.9	192	9	US-09-738-626-6797
31	46	46.9	265	16	US-10-389-566-2422
32	46	46.9	533	9	US-09-808-602-87
33	46	46.9	533	10	US-09-808-198-74
34	45.5	46.4	318	9	US-09-738-626-4910
35	45	45.9	181	13	US-10-062-254-369
36	45	45.9	317	14	US-10-080-170-167
37	45	45.9	400	14	US-10-301-822-155
38	45	45.9	492	14	US-10-176-306-55
39	45	45.9	564	13	US-10-139-262-28
40	45	45.9	564	14	US-10-255-969-28
41	44.5	45.4	200	15	US-10-264-237-2258
42	44.5	45.4	241	14	US-10-106-698-4882
43	44.5	45.4	293	14	US-10-012-542-233
44	44.5	45.4	307	14	US-10-012-542-331
45	44.5	45.4	434	11	US-09-764-875-783

ALIGNMENTS

RESULT 1

US-10-162-335-76
; Sequence 76, Application US/10162335
; Publication No. US20040009480A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, David W.
; APPLICANT: Baumgartner, Jason C.
; APPLICANT: Boldog, Ferenc L.
; APPLICANT: Casman, Stacie J.
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Gangolli, Esha A.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Gorman, Linda
; APPLICANT: Guo, Xiaojia (Sasha)
; APPLICANT: Hjalte, Tord
; APPLICANT: Kexuda, Ramesh
; APPLICANT: Li, Li
; APPLICANT: MacDougall, John R.
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Millet, Isabelle
; APPLICANT: Padigar, Maralidara
; APPLICANT: Paturajan, Meera
; APPLICANT: Pena, Carol E. A.
; APPLICANT: Rastelli, Luca
; APPLICANT: Shimkets, Richard A.
; APPLICANT: Stone, David J.
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Vernet, Corine A. M.
; APPLICANT: Voss, Edward Z.
; APPLICANT: Zertusen, Bryan D.
; TITLE OF INVENTION: Therapeutic Polypeptides, Nucleic Acids Encoding Same, and Metho
; FILE REFERENCE: 21402-377 B
; CURRENT APPLICATION NUMBER: US/10162,335
; CURRENT FILING DATE: 2003-10-01
; PRIOR APPLICATION NUMBER: 60/295,607
; PRIOR FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: 60/295,661
; PRIOR FILING DATE: 2001-06-04

Sequence 2153, Ap
Sequence 22, Appl
Sequence 20, Appl
Sequence 24, Appl
Sequence 8065, Ap
Sequence 6, Appl
Sequence 66, Appl
Sequence 30, Appl
Sequence 6826, Ap
Sequence 170, App
Sequence 797, App
Sequence 33, Appl
Sequence 2423, Ap
Sequence 2424, Ap
Sequence 6797, Ap
Sequence 2422, Ap
Sequence 87, Appl
Sequence 74, Appl
Sequence 4310, Ap
Sequence 369, App
Sequence 167, App
Sequence 155, App
Sequence 55, App
Sequence 28, Appl
Sequence 28, Appl
Sequence 2256, Ap
Sequence 4882, Ap
Sequence 233, App
Sequence 331, App
Sequence 783, App

```

; PRIOR APPLICATION NUMBER: 60/296,404
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: 60/296,418
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: 60/297,414
; PRIOR FILING DATE: 2001-06-11
; PRIOR APPLICATION NUMBER: 60/297,567
; PRIOR FILING DATE: 2001-06-12
; PRIOR APPLICATION NUMBER: 60/298,285
; PRIOR FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: 60/298,556
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/299,949
; PRIOR FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: 60/300,883
; PRIOR FILING DATE: 2001-06-26
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 201
; SEQ ID NO 76
; LENGTH: 415
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-162-335-76

```

```

Query Match      98.0%; Score 96; DB 15; Length 415;
Best Local Similarity 94.1%; Pred. No. 8.1e-06;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 HGHEQOQHLGHGHKXKL 17
|||||
DB 234 HGHEQOQHLGHGHKFKL 250

```

```

RESULT 2
US-10-162-335-72
; Sequence 72, Application US/101622335
; Publication No. US20040009480A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, David W.
; APPLICANT: Baumgartner, Jason C.
; APPLICANT: Boldog, Ferenc L.
; APPLICANT: Casman, Stacie J.
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Gangolli, Esha A.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Gorman, Linda
; APPLICANT: Guo, Xiaojia (Sasha)
; APPLICANT: Hjalt, Tord
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Li, Li
; APPLICANT: MacDougall, John R.
; APPLICANT: Malvankar, Uriel M.
; APPLICANT: Millet, Isabelle
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Patturajan, Meera
; APPLICANT: Pena, Carol E. A.
; APPLICANT: Rastelli, Luca
; APPLICANT: Shimkets, Richard A.
; APPLICANT: Stone, David J.
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Vernet, Corine A. M.
; APPLICANT: Voss, Edward Z.
; APPLICANT: Zernhusen, Bryan D.
; TITLE OF INVENTION: Therapeutic Polypeptides, Nucleic Acids Encoding Same, and Method
; FILE REFERENCE: 21402-377 B
; CURRENT APPLICATION NUMBER: US/10/162,335
; CURRENT FILING DATE: 2002-10-01
; PRIOR APPLICATION NUMBER: 60/295,607
; PRIOR FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: 60/295,661
; PRIOR FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: 60/296,404
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: 60/296,418
; PRIOR FILING DATE: 2001-06-06

```

```

; PRIOR APPLICATION NUMBER: 60/296,418
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: 60/297,414
; PRIOR FILING DATE: 2001-06-11
; PRIOR APPLICATION NUMBER: 60/297,567
; PRIOR FILING DATE: 2001-06-12
; PRIOR APPLICATION NUMBER: 60/298,285
; PRIOR FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: 60/298,556
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/299,949
; PRIOR FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: 60/300,883
; PRIOR FILING DATE: 2001-06-26
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 201
; SEQ ID NO 72
; LENGTH: 615
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-162-335-72

```

```

Query Match      98.0%; Score 96; DB 15; Length 615;
Best Local Similarity 94.1%; Pred. No. 1.2e-05;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 HGHEQOQHLGHGHKXKL 17
|||||
DB 434 HGHEQOQHLGHGHKFKL 450

```

```

RESULT 3
US-10-162-335-74
; Sequence 74, Application US/10162335
; Publication No. US20040009480A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, David W.
; APPLICANT: Baumgartner, Jason C.
; APPLICANT: Boldog, Ferenc L.
; APPLICANT: Casman, Stacie J.
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Gangolli, Esha A.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Gorman, Linda
; APPLICANT: Guo, Xiaojia (Sasha)
; APPLICANT: Hjalt, Tord
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Li, Li
; APPLICANT: MacDougall, John R.
; APPLICANT: Malvankar, Uriel M.
; APPLICANT: Millet, Isabelle
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Patturajan, Meera
; APPLICANT: Pena, Carol E. A.
; APPLICANT: Rastelli, Luca
; APPLICANT: Shimkets, Richard A.
; APPLICANT: Stone, David J.
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Vernet, Corine A. M.
; APPLICANT: Voss, Edward Z.
; APPLICANT: Zernhusen, Bryan D.
; TITLE OF INVENTION: Therapeutic Polypeptides, Nucleic Acids Encoding Same, and Method
; FILE REFERENCE: 21402-377 B
; CURRENT APPLICATION NUMBER: US/10/162,335
; CURRENT FILING DATE: 2002-10-01
; PRIOR APPLICATION NUMBER: 60/295,607
; PRIOR FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: 60/295,661
; PRIOR FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: 60/296,404
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: 60/296,418
; PRIOR FILING DATE: 2001-06-06

```

seq-a.rapb

Fri Mar 5 08:53:28 2004

```

; PRIOR APPLICATION NUMBER: 60/297,567
; PRIOR FILING DATE: 2001-06-12
; PRIOR APPLICATION NUMBER: 60/298,285
; PRIOR FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: 60/298,556
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/299,949
; PRIOR FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: 60/300,883
; PRIOR FILING DATE: 2001-06-26
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 201
; SEQ ID NO 84
; LENGTH: 644
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-162-335-84

Query Match      98.0%; Score 96; DB 15; Length 644;
Best Local Similarity 94.1%; Pred. No. 1.2e-05;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 HGHEQQHGLGHGKXKL 17
      |||||
DB      463 HGHEQQHGLGHGKFKL 479

RESULT 5
US-10-156-761-14106
; Sequence 14106, Application US/10:56761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASHIEA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 14106
; LENGTH: 574
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-14106

Query Match      61.2%; Score 60; DB 14; Length 574;
Best Local Similarity 64.3%; Pred. No. 1.3;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY      1 HGHEQQHGLGHGK 14
      |||||
DB      418 HKHNEHGHGHGK 431

RESULT 6
US-10-270-333-42
; Sequence 42, Application US/10:270333
; Publication No. US20030092124A1
; GENERAL INFORMATION:
; APPLICANT: Cravchik, Anibal
; TITLE OF INVENTION: ISOLATED G-PROTEIN COUPLED RECEPTORS.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING GPCR PROTEINS, AND USES
; TITLE OF INVENTION: THEROP AS INSECTICIDAL TARGETS
; FILE REFERENCE: CL000733CON

```

seq-a.rapb

Fri Mar 5 08:53:28 2004

```
; NUMBER OF SEQ ID NOS: 86
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 29
; LENGTH: 376
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-176-306-29

Query Match      56.1%; Score 55; DB 14; Length 376;
Best Local Similarity 56.2%; Pred. No. 4.3; 7; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy      2 GHEQQHGLGHGHKXKL 17
Db      163 GHGSHGSGHGHSHSL 178

RESULT 8
US-10-264-237-1632
; Sequence 1632, Application US/10264237
; Publication No. US20040009491A1
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA131P1
; CURRENT APPLICATION NUMBER: US/10/264,237
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/16450
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: US 60/205,515
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 2876
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 1632
; LENGTH: 376
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (15)-
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (59)-
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (149)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (364)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
US-10-264-237-1632

Query Match      56.1%; Score 55; DB 15; Length 376;
Best Local Similarity 56.2%; Pred. No. 4.3;
Matches 9; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy      2 GHEQQHGLGHGHKXKL 17
Db      163 GHGSHGSGHGHSHSL 178

RESULT 9
US-10-176-306-59
; Sequence 59, Application US/10176306
; Publication No. US20030130485A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel A.
; APPLICANT: Curtis, Rory A. J.
; APPLICANT: Glucksmann, Maria Alexandra
; APPLICANT: Bandaru, Rajasekhar
```


seq-a.rapb

Fri Mar 5 08:53:28 2004

```
; ORGANISM: Mus musculus
US-09-738-897-5
Query Match      55.1%; Score 54; DB 9; Length 503;
Best Local Similarity 60.0%; Pred. No. 7.9;
Matches 9; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY      2 GHEQOQHGLGHGKXK 16
Db      145 GHGSHSGHGHGLAK 159

RESULT 11
US-09-738-897-4
; Sequence 4, Application US/09738897
; Patent No. US20020106721A1
; GENERAL INFORMATION:
; APPLICANT: GUEGLER, Karl et al
; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL001053
; CURRENT APPLICATION NUMBER: US/09/738,897
; CURRENT FILING DATE: 2000-12-18
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 507
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-738-897-4
Query Match      55.1%; Score 54; DB 9; Length 507;
Best Local Similarity 60.0%; Pred. No. 8;
Matches 9; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY      2 GHEQOQHGLGHGKXK 16
Db      145 GHGSHSGHGHGLAK 159

RESULT 12
US-09-849-138-41
; Sequence 41, Application US/09849138
; Publication No. US20030139358A1
; GENERAL INFORMATION:
; APPLICANT: Spytek, Kimberly
; APPLICANT: Padigaru, Maralidhara
; APPLICANT: Majumder, Kumud
; APPLICANT: MacDougall, John
; APPLICANT: Stone, David
; APPLICANT: Gangolli, Bsha
; APPLICANT: Spaderna, Glennda
; APPLICANT: Smithson, Glennda
; TITLE OF INVENTION: No. US20030139358A1el Human Proteins, Polynucleotides Encoding Ti
; TITLE OF INVENTION: Methods Using the Same
; FILE REFERENCE: 21402-001 US
; CURRENT APPLICATION NUMBER: US/05/849,138
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: 60/201,951
; PRIOR FILING DATE: 2000-05-20
; PRIOR APPLICATION NUMBER: 60/215,857
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/265,162
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/203,109
; PRIOR FILING DATE: 2000-05-08
; PRIOR APPLICATION NUMBER: 60/203,295
; PRIOR FILING DATE: 2000-05-11
; PRIOR APPLICATION NUMBER: 60/210,055
; PRIOR FILING DATE: 2000-06-07
; PRIOR APPLICATION NUMBER: 60/204,064
; PRIOR FILING DATE: 2000-05-12

; APPLICANT: Kapeller-Libermann, Rosana
; TITLE OF INVENTION: NOVEL HUMAN GENES AND METHODS OF USE THEREOF
; FILE REFERENCE: 10448-195001
; CURRENT APPLICATION NUMBER: US/10/176,306
; CURRENT FILING DATE: 2002-06-20
; PRIOR APPLICATION NUMBER: 10/001,137
; PRIOR FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: PCT/US01/45291
; PRIOR FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/248,362
; PRIOR FILING DATE: 2000-11-14
; PRIOR APPLICATION NUMBER: 60/248,331
; PRIOR FILING DATE: 2000-11-14
; PRIOR APPLICATION NUMBER: 60/248,365
; PRIOR FILING DATE: 2000-11-14
; PRIOR APPLICATION NUMBER: 60/250,077
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: 60/250,327
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: 60/250,176
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: 10/023,617
; PRIOR FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: PCT/US01/49416
; PRIOR FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: 60/256,249
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: 60/256,405
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: 10/083,248
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: PCT/US01/46717
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: 60/242,324
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/242,518
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/241,989
; PRIOR FILING DATE: 2000-10-20
; NUMBER OF SEQ ID NOS: 86
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 59
; LENGTH: 378
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-176-306-59
Query Match      56.1%; Score 55; DB 14; Length 378;
Best Local Similarity 56.2%; Pred. No. 4.4;
Matches 9; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY      2 GHEQOQHGLGHGKXKL 17
Db      163 GHGSHSGHGHGHSLS 178

RESULT 10
US-09-738-897-5
; Sequence 5, Application US/09738897
; Patent No. US20020106721A1
; GENERAL INFORMATION:
; APPLICANT: GUEGLER, Karl et al
; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL001053
; CURRENT APPLICATION NUMBER: US/09/738,897
; CURRENT FILING DATE: 2000-12-18
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 503
; TYPE: PRT
```

; PRIOR APPLICATION NUMBER: 60/204,063
; PRIOR FILING DATE: 2000-05-12
; PRIOR APPLICATION NUMBER: 60/204,062
; PRIOR FILING DATE: 2000-05-12
; PRIOR APPLICATION NUMBER: 60/203,838
; PRIOR FILING DATE: 2000-05-12
; PRIOR APPLICATION NUMBER: 60/203,839
; PRIOR FILING DATE: 2000-05-12
; PRIOR APPLICATION NUMBER: 60/204,089
; PRIOR FILING DATE: 2000-05-15
; PRIOR APPLICATION NUMBER: 60/204,276
; PRIOR FILING DATE: 2000-05-16
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 41
; LENGTH: 507
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-849-138-41

Query Match 55.1%; Score 54; DB 10; Length 507;
Best Local Similarity 60.0%; Pred. No. 8;
Matches 9; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 GHEQOQHGLGHGXK 16
DB 145 GHGSHGHGHGLAK 159

RESULT 13
US-10-176-306-49
; Sequence 49, Application US/10176306
; Publication No. US20030130495A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel A.
; APPLICANT: Curtis, Rory A. J.
; APPLICANT: Glucksmann, Maria Alexandra
; APPLICANT: Bandaru, Rajasekhar
; APPLICANT: Kapeller-Liebermann, Rosana
; TITLE OF INVENTION: NOVEL HUMAN GENES AND METHODS OF USE THEREOF
; FILE REFERENCE: 10448-195001
; CURRENT APPLICATION NUMBER: US/10/176,306
; CURRENT FILING DATE: 2002-06-20
; PRIOR APPLICATION NUMBER: 10/001,137
; PRIOR FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: PCT/US01/45291
; PRIOR FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/248,362
; PRIOR FILING DATE: 2000-11-14
; PRIOR APPLICATION NUMBER: 60/248,331
; PRIOR FILING DATE: 2000-11-14
; PRIOR APPLICATION NUMBER: 60/248,365
; PRIOR FILING DATE: 2000-11-14
; PRIOR APPLICATION NUMBER: 60/250,077
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: 60/250,327
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: 60/250,176
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: 10/023,617
; PRIOR FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: PCT/US01/49416
; PRIOR FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: 60/256,249
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: 60/256,405
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: 10/083,248
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: PCT/US01/46717
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: 60/242,324
; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/242,518
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/241,989
; PRIOR FILING DATE: 2000-10-20
; NUMBER OF SEQ ID NOS: 86
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 49
; LENGTH: 507
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-10-176-306-49

Query Match 55.1%; Score 54; DB 14; Length 507;
Best Local Similarity 60.0%; Pred. No. 8;
Matches 9; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 GHEQOQHGLGHGXK 16
DB 145 GHGSHGHGHGLAK 159

RESULT 14
US-10-360-849A-36
; Sequence 36, Application US/10360849A
; Publication No. US20030220249A1
; GENERAL INFORMATION:
; APPLICANT: Discovery Genomics, Inc.
; APPLICANT: Rackett, Perry
; APPLICANT: Nasevicius, Aidan
; APPLICANT: Bessner, Jeffrey
; APPLICANT: Clark, Karl
; APPLICANT: Larson, Jon
; APPLICANT: Ekker, Stephen
; APPLICANT: Roberg-Perez, Sharon
; APPLICANT: Madman, Shannon
; TITLE OF INVENTION: FACTORS FOR ANGIOGENESIS, VASCULOGENESIS, CARTILAGE FORMATION,
; FILE REFERENCE: 3021.05US02
; CURRENT APPLICATION NUMBER: US/10/360,849A
; CURRENT FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: US 60/354,978
; PRIOR FILING DATE: 2002-02-07
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 36
; LENGTH: 469
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-360-849A-36

Query Match 54.1%; Score 53; DB 15; Length 469;
Best Local Similarity 57.1%; Pred. No. 10;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 HGHEQOQHGLGHGXK 14
DB 56 HGHSHGHGHGHGTHE 69

RESULT 15
US-09-789-561-86
; Sequence 86, Application US/09789561
; Patent No. US20020064818A1
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: S2 Human secreted proteins
; FILE REFERENCE: PZ043P1
; CURRENT APPLICATION NUMBER: US/09/789,561
; CURRENT FILING DATE: 2001-02-22
; PRIOR APPLICATION NUMBER: PCT/US00/24008
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: 60/152,317
; PRIOR FILING DATE: 1999-09-03

seq-a.rapb

Fri Mar 5 08:53:28 2004

```

; PRIOR APPLICATION NUMBER: 60/152,315
; PRIOR FILING DATE: 1999-09-03
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: Patentm Ver. 2.0
; SEQ ID NO 86
; LENGTH: 831
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-789-561-86

Query Match      54.1%; Score 53; DB 9; Length 831;
Best Local Similarity 61.5%; Pred. No. 18;
Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY      1 HGHEQQGGLGHGH 13
DB      467 HDHSHQHAGHGH 479

Search completed: March 4, 2004, 13:26:32
Job time : 26 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2004, 12:52:17 ; Search time 47.25 Seconds
(without alignments)
101.657 Million cell updates/sec

Title: SEQ-A

Perfect score: 98

Sequence: 1 HGHEQQHGLGHGKXK 17

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_29Jan04.*

- 1: geneseqp1990s.*
- 2: geneseqp1990s.*
- 3: geneseqp2000s.*
- 4: geneseqp2001s.*
- 5: geneseqp2002s.*
- 6: geneseqp2003as.*
- 7: geneseqp2003bs.*
- 8: geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	96	98.0	28	3	AAy81996 Human hig
2	96	98.0	55	3	AAy93346 Light cha
3	96	98.0	62	3	AAy93348 Light cha
4	96	98.0	63	2	AAr75186 Partial p
5	96	98.0	83	3	AAy93347 Light cha
6	96	98.0	94	3	AAy93351 Light cha
7	96	98.0	125	5	ABb78708 Human hig
8	96	98.0	131	2	AAr75181 Partial p
9	96	98.0	158	5	ABb78709 Calmoduli
10	96	98.0	186	3	AAy93349 Light cha
11	96	98.0	255	3	AAy93342 Light cha
12	96	98.0	255	4	ABg73620 Human hig
13	96	98.0	305	4	ABg21100 Novel hum
14	96	98.0	357	6	ABr41202 Human DIT
15	96	98.0	415	6	ABu99146 Novel hum
16	96	98.0	615	6	ABu99144 Novel hum
17	96	98.0	626	5	ABb78707 Human hig
18	96	98.0	644	4	ABg21101 Novel hum
19	96	98.0	644	5	ABb78710 Human hig
20	96	98.0	644	5	ABu99150 Novel hum
21	96	98.0	644	6	ABu99145 Novel hum
22	96	98.0	720	4	ABg21103 Novel hum
23	88	89.8	17	2	AAw07627 Human hig
24	81	82.7	41	2	AAr75180 Partial p
25	81	82.7	110	2	AAr75178 Partial p

ALIGNMENTS

RESULT 1

AAy81996
ID AAy81996 standard; peptide; 28 AA.

XX AC AAy81996;

XX 16-OCT-2000 (first entry)

XX Human high molecular weight kininogen domain 5 fragment #5.

XX Human; high molecular weight kininogen; HK;
KW two-chain high molecular weight kininogen; HKa; angiogenesis inhibition;
KW tumour; cancer; ocular disorder; rheumatoid arthritis;
KW endothelial cell apoptosis.

XX Homo sapiens.

XX WO200027866-A1.

XX 18-MAY-2000.

XX 05-NOV-1999; 99WO-US026419.

XX 10-NOV-1998; 98US-0107833P.

XX (UTEM) UNIV TEMPLE.

XX (MCCR/) MCCRAE R K.

XX Mccrae RK;

XX WPI; 2000-376483/32.

XX A pharmaceutical composition used to inhibit angiogenesis, inhibit endothelial cell proliferation, and induce endothelial cell apoptosis.

XX Claim 8; Page 28; 52pp; English.

XX The present sequence is derived from human high molecular weight kininogen (HK) domain 5. HK is a 120 kD glycoprotein which binds with high affinity to endothelial cells, where it is cleaved to two-chain high molecular weight kininogen (HKa) by plasma kallikrein. HKa or a synthetic compound comprising the present sequence may be used in a pharmaceutical composition for inhibiting angiogenesis. Angiogenesis occurs in a number of disease states, such as tumour formation and expansion, and certain ocular disorders. It can also occur in a rheumatoid joint, hastening joint destruction by allowing an influx of leukocytes. The composition may inhibit angiogenesis by inhibiting endothelial cell proliferation or by inducing endothelial cell apoptosis. Peptides used in the composition

had date

CC may be recombinant peptides, natural peptides, or synthetic peptides.
 CC They may also be chemically synthesised, using, for example, solid phase
 CC synthesis methods

XX SQ Sequence 28 AA;

Query Match 98.0%; Score 96; DB 3; Length 28;
 Best Local Similarity 94.1%; Pred. No. 2e-07;
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 HGHEQOQHLGHGKXKL 17
 Db 1 HGHEQOQHLGHGKXKL 17

RESULT 2
 AAY93346
 ID AAY93346 standard; peptide; 55 AA.

XX AC AAY93346;
 XX DT 04-SEP-2000 (first entry)

XX DE Light chain of human high molecular weight kininogen analogue.

XX KW Human; high molecular weight kininogen; glycoprotein; endothelial cell;
 KW plasma kallikrein; heavy chain; light chain; analogue; angiogenesis;
 KW endothelial cell proliferation; endothelial cell migration; vitronectin.

XX OS Synthetic.
 XX OS Homo sapiens.

XX PN WO200027415-A2.

XX PD 18-MAY-2000.

XX PF 09-NOV-1999; 99WO-US026377.

XX PR 10-NOV-1998; 98US-010784P.

XX PA (UTEM) UNIV TEMPLE.

XX PA (DUPO) DUPONT PHARM CO.

XX PA (COLM/) COLMAN W R.

XX PA (MOUS/) MOUSA A S.

XX PI Colman WR, Mousa AS;

XX DR WPI; 2000-376306/32.

XX PT Method for inhibiting endothelial cell proliferation, using compound that
 PT inhibit endothelial cell migration.

XX PS Claim 4; Page 36; 41pp; English.

XX CC The present sequence represents an analogue of the light chain of human
 CC high molecular weight kininogen. High molecular weight kininogen is a 120
 CC kDa glycoprotein which binds with high affinity to endothelial cells,
 CC where it is cleaved by plasma kallikrein into heavy and light chains.
 CC Analogues of high molecular weight kininogen are used in the method of
 CC the invention. The specification describes a method of inhibiting
 CC endothelial cell proliferation. The method comprises contacting
 CC endothelial cells with a compound containing high molecular weight
 CC kininogen analogues. The method and the compounds can be used for
 CC inhibiting endothelial cell proliferation. The compounds can also be used
 CC for inhibiting angiogenesis. The compounds can also be used to inhibit
 CC migration of endothelial cells to vitronectin

XX SQ Sequence 55 AA;

Query Match 98.0%; Score 96; DB 3; Length 55;
 Best Local Similarity 94.1%; Pred. No. 4.2e-07;
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 HGHEQOQHLGHGKXKL 17
 Db 26 HGHEQOQHLGHGKXKL 42

RESULT 3
 AAY93348
 ID AAY93348 standard; peptide; 62 AA.

XX AC AAY93348;

XX DT 04-SEP-2000 (first entry)

XX DE Light chain of human high molecular weight kininogen analogue.

XX KW Human; high molecular weight kininogen; glycoprotein; endothelial cell;
 KW plasma kallikrein; heavy chain; light chain; analogue; angiogenesis;
 KW endothelial cell proliferation; endothelial cell migration; vitronectin.

XX OS Synthetic.

XX OS Homo sapiens.

XX PN WO200027415-A2.

XX PD 18-MAY-2000.

XX PF 09-NOV-1999; 99WO-US026377.

XX PR 10-NOV-1998; 98US-010784P.

XX PA (UTEM) UNIV TEMPLE.

XX PA (DUPO) DUPONT PHARM CO.

XX PA (COLM/) COLMAN W R.

XX PA (MOUS/) MOUSA A S.

XX PI Colman WR, Mousa AS;

XX DR WPI; 2000-376306/32.

XX PT Method for inhibiting endothelial cell proliferation, using compound that
 PT inhibit endothelial cell migration.

XX PS Claim 6; Page 37; 41pp; English.

XX CC The present sequence represents an analogue of the light chain of human
 CC high molecular weight kininogen. High molecular weight kininogen is a 120
 CC kDa glycoprotein which binds with high affinity to endothelial cells,
 CC where it is cleaved by plasma kallikrein into heavy and light chains.
 CC Analogues of high molecular weight kininogen are used in the method of
 CC the invention. The specification describes a method of inhibiting
 CC endothelial cell proliferation. The method comprises contacting
 CC endothelial cells with a compound containing high molecular weight
 CC kininogen analogues. The method and the compounds can be used for
 CC inhibiting endothelial cell proliferation. The compounds can also be used
 CC for inhibiting angiogenesis. The compounds can also be used to inhibit
 CC migration of endothelial cells to vitronectin

XX SQ Sequence 62 AA;

Query Match 98.0%; Score 96; DB 3; Length 62;
 Best Local Similarity 94.1%; Pred. No. 4.7e-07;
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 HGHEQOQHLGHGKXKL 17
 Db 5 HGHEQOQHLGHGKXKL 21

RESULT 4
 AAY75186
 ID AAY75186 standard; peptide; 63 AA.

XX AC AAY75186;

XX DT 05-DEC-1995 (first entry)
 XX DE Partial peptide of human HMW kininogen fragment 2.
 XX KW high molecular weight; kininogen; fragment; 1.2; 1; 2; partial;
 XX KW wound treating agent; bovine; growth promotion; fibroblast.
 XX OS Homo sapiens.
 XX FN JP07082172-A.
 XX PD 28-MAR-1995.
 XX PF 17-SEP-1993; 93JP-00230616.
 XX PR 17-SEP-1993; 93JP-00230616.
 XX PA (PARH) HOECHST JAPAN KK.
 XX DR WPI; 1995-158909/21.
 XX PT A wound treating agent contg. a partial peptide of kininogen - have
 XX PT growth promotion activity of fibroblasts.
 XX PS Claim 8; Page 8; 8pp; Japanese.
 XX CC AAR75186 is a partial peptide corresponding to human kininogen fragment
 CC 1, amino acids 458-520. Partial peptides of bovine and human kininogen
 CC fragments 1.2, 1 and 2, are used in wound treating agent compns. and act
 CC as the active component. The fragments are useful in wound treating
 CC because they have growth promotion activity on fibroblasts
 XX SQ Sequence 63 AA;
 XX Query Match 98.0%; Score 96; DB 2; Length 63;
 XX Best Local Similarity 94.1%; Pred. No. 4.8e-07;
 XX Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 HGHEQQHGLGHGKFKL 17
 DB |||||
 6 HGHEQQHGLGHGKFKL 22
 RESULT 5
 AAY93347
 ID AAY93347 standard; peptide; 83 AA.
 XX AC AAY93347;
 XX DT 04-SEP-2000 (first entry)
 XX DE Light chain of human high molecular weight kininogen analogue.
 XX KW Human; high molecular weight kininogen; glycoprotein; endothelial cell;
 XX KW plasma kallikrein; heavy chain; light chain; analogue; angiogenesis;
 XX KW endothelial cell proliferation; endothelial cell migration; vitronectin.
 XX OS Synthetic.
 XX OS Homo sapiens.
 XX FN WC200027415-A2.
 XX PD 18-MAY-2000.
 XX PF 09-NOV-1999; 99WO-US026377.
 XX PR 10-NOV-1998; 98US-0107844P.
 XX PA (UTEM) UNIV TEMPLE.
 XX PA (DUPO) DUPONT PHARM CO.
 XX PA (COLM/) COLMAN W R.
 XX PA (MOUS/) MOUSA A S.
 XX PI Colman WR, Mousa AS;
 XX DR WPI; 2000-376306/32.
 XX PT Method for inhibiting endothelial cell proliferation, using compound that
 XX PT inhibit endothelial cell migration.
 XX PS Claim 8; Page 39; 41pp; English.

XX PI Colman WR, Mousa AS;
 XX DR WPI; 2000-376306/32.
 XX KW Method for inhibiting endothelial cell proliferation, using compound that
 XX PT inhibit endothelial cell migration.
 XX PS Claim 5; Page 37; 41pp; English.
 XX CC The present sequence represents an analogue of the light chain of human
 CC high molecular weight kininogen. High molecular weight kininogen is a 120
 CC kDa glycoprotein which binds with high affinity to endothelial cells,
 CC where it is cleaved by plasma kallikrein into heavy and light chains.
 CC Analogues of high molecular weight kininogen are used in the method of
 CC the invention. The specification describes a method of inhibiting
 CC endothelial cell proliferation. The method comprises contacting
 CC endothelial cells with a compound containing high molecular weight
 CC kininogen analogues. The method and the compounds can be used for
 CC inhibiting endothelial cell proliferation. The compounds can also be used
 CC for inhibiting angiogenesis. The compounds can also be used to inhibit
 CC migration of endothelial cells to vitronectin
 XX SQ Sequence 83 AA;
 XX Query Match 98.0%; Score 96; DB 3; Length 83;
 XX Best Local Similarity 94.1%; Pred. No. 6.4e-07;
 XX Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 HGHEQQHGLGHGKFKL 17
 DB |||||
 26 HGHEQQHGLGHGKFKL 42
 RESULT 6
 AAY93351
 ID AAY93351 standard; peptide; 94 AA.
 XX AC AAY93351;
 XX DT 04-SEP-2000 (first entry)
 XX DE Light chain of human high molecular weight kininogen analogue.
 XX KW Human; high molecular weight kininogen; glycoprotein; endothelial cell;
 XX KW plasma kallikrein; heavy chain; light chain; analogue; angiogenesis;
 XX KW endothelial cell proliferation; endothelial cell migration; vitronectin.
 XX OS Synthetic.
 XX OS Homo sapiens.
 XX FN WC200027415-A2.
 XX PD 18-MAY-2000.
 XX PF 09-NOV-1999; 99WO-US026377.
 XX PR 10-NOV-1998; 98US-0107844P.
 XX PA (UTEM) UNIV TEMPLE.
 XX PA (DUPO) DUPONT PHARM CO.
 XX PA (COLM/) COLMAN W R.
 XX PA (MOUS/) MOUSA A S.
 XX PI Colman WR, Mousa AS;
 XX DR WPI; 2000-376306/32.
 XX PT Method for inhibiting endothelial cell proliferation, using compound that
 XX PT inhibit endothelial cell migration.
 XX PS Claim 8; Page 39; 41pp; English.

CC The present sequence represents an analogue of the light chain of human
 CC high molecular weight kininogen. High molecular weight kininogen is a 120
 CC kDa glycoprotein which binds with high affinity to endothelial cells,
 CC where it is cleaved by plasma kallikrein into heavy and light chains.
 CC Analogues of high molecular weight kininogen are used in the method of
 CC the invention. The specification describes a method of inhibiting
 CC endothelial cell proliferation. The method comprises contacting
 CC endothelial cells with a compound containing high molecular weight
 CC kininogen analogues. The method and the compounds can be used for
 CC inhibiting endothelial cell proliferation. The compounds can also be used
 CC for inhibiting angiogenesis. The compounds can also be used to inhibit
 CC migration of endothelial cells to vitronectin
 XX

SQ Sequence 94 AA;

Query Match 98.0%; Score 96; DB 3; Length 94;
 Best Local Similarity 94.1%; Pred. No. 7.3e-07;
 Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 HGHEQHQHGLGHGKXKL 17
 |||||
 Db 26 HGHEQHQHGLGHGKFKL 42
 |||||

RESULT 7

ABB78708
 ID ABB78708 standard; protein; 125 AA.

AC ABB78708;

XX 18-JUL-2002 (first entry)

XX Human high molecular weight kininogen (HK) D5 domain SEQ ID NO:2.

XX Human: kininogen; high molecular weight kininogen; HK; D5 domain;
 KW D5 receptor; angiogenesis; endothelial cell; cytostatic; antitumor;
 KW antithrombotic; vasectopic; vulnary; tranquiliser; thrombolytic;
 KW ophthalmological; gynaecological; antitumor; antidiabetic; antithrombotic;
 XX antiangiogenic; antiapoptotic; endocrine; apoptosis; gene therapy.

XX Homo sapiens.

XX WO200214369-A2.

XX 21-FEB-2002.

XX 24-JUL-2001; 2001WO-US023185.

XX 24-JUL-2000; 2000US-0220194P.

XX (ATTE-) ATTENUON LLC.

XX Mazar AP, Juarez JC;

XX WPI; 2002-393611/42.

XX N-PSDE; ASL52521.

XX Novel human kininogen D5 domain polypeptides useful for treating
 PT conditions associated with endothelial cell migration, proliferation,
 PT invasion or angiogenesis, e.g. arthritis, macular degeneration, benign
 PT hyperplasia.

XX Claim 2; Page 13; 84pp; English.

XX The present invention describes an isolated polypeptide (I) that
 CC corresponds to the D5 domain of human kininogen, or biologically active
 CC peptide fragment, homologue or functional derivative, and which: (a)
 CC inhibits angiogenesis; (b) binds to the D5 binding site on endothelial
 CC cells [EC]; (c) activates signalling pathways leading to the introduction
 CC of apoptosis in EC; and/or (d) inhibits the signalling pathway required
 CC for maintenance of EC viability. (I) has cytostatic, antitumor,
 CC antithrombotic, vasectopic, vulnary, tranquiliser, thrombolytic,
 CC ophthalmological, gynaecological, antitumor, antidiabetic, antithrombotic,

CC antiangiogenic, antiapoptotic and endocrine activities. An antibody (IX)
 CC specific for an epitope of (I) is useful for inhibiting tumour growth or
 CC angiogenesis in a subject. (I), a D5 fusion polypeptide (II) or a dimeric
 CC or trimeric fusion polypeptide (III) can be used for inhibiting EC
 CC migration, proliferation, invasion, or angiogenesis, or for inducing EC
 CC apoptosis. An angiogenic EC-targeting pharmaceutical composition (X)
 CC comprising (I), (II), or (III), can be used for treating a subject having
 CC a disease or condition associated with undesired EC migration,
 CC proliferation, invasion or angiogenesis. (I), (II), or (III) can be used
 CC for isolating a D5 domain binding molecule from a complex mixture and for
 CC isolating or enriching cells expressing D5 domain binding sites from a
 CC cell mixture. The present sequence represents the human high molecular
 CC weight kininogen (HK) D5 domain amino acid sequence, which is given in
 CC the exemplification of the present invention
 XX

SQ Sequence 125 AA;

Query Match 98.0%; Score 96; DB 5; Length 125;
 Best Local Similarity 94.1%; Pred. No. 9.9e-07;
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 HGHEQHQHGLGHGKXKL 17
 |||||

Db 62 HGHEQHQHGLGHGKFKL 78
 |||||

RESULT 8

AAR75181

ID AAR75181 standard; peptide; 131 AA.

XX AAR75181;

XX 05-DEC-1995 (first entry)

XX Partial peptide of human HKW kininogen fragment 1.2.

XX high molecular weight; kininogen; fragment; 1.2; 1; 2; partial;
 KW wound treating agent; human; growth promotion; fibroblast.

XX Homo sapiens.

XX JP07082172-A.

XX 28-MAR-1995.

XX 17-SEP-1993; 93JP-00230616.

XX 17-SEP-1993; 93JP-00230616.

XX (PARH) HOECHST JAPAN KK.

XX WPI; 1995-158909/21.

XX A wound treating agent contg. a partial peptide of kininogen - have
 PT growth promotion activity of fibroblasts.

XX Claim 7; Page 7; 8pp; Japanese.

XX AAR75181 is a partial peptide corresponding to human kininogen fragment
 CC 1.2, amino acids 350-520. Partial peptides of bovine and human kininogen
 CC fragments 1.2, 1 and 2, are used in wound treating agent compns. and act
 CC as the active component. The fragments are useful in wound treating
 CC because they have growth promotion activity on fibroblasts
 XX

SQ Sequence 131 AA;

Query Match 98.0%; Score 96; DB 2; Length 131;
 Best Local Similarity 94.1%; Pred. No. 1e-06;
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 HGHEQHQHGLGHGKXKL 17
 |||||

Db 74 HGHEQHQHGLGHGKFKL 90
 |||||

```

RESULT 9
ABE78709
ID ABE78709 standard; protein; 158 AA.
XX
AC ABE78709;
XX
DT 18-JUL-2002 (first entry)
XX
DE Calmodulin binding protein and D5 domain fusion protein SEQ ID NO:13.
XX
KW Human; kininogen; high molecular weight kininogen; HK; D5 domain;
KW D5 receptor; angiotensin; endothelial cell; cytosolic; antitumor;
KW antithrombotic; vasotrophic; vulnar; tranquilliser; thrombolytic;
KW ophthalmological; gynaecological; antiulcer; antidiabetic; antiarthritic;
KW antiangiogenic; antipapillary; endocrine; apoptosis; gene therapy;
KW calmodulin binding protein; CBP.
XX
OS Homo sapiens.
XX
PH Key Location/Qualifiers
FT Domain 1..29
FT /note= "Calmodulin binding protein (CBP) sequence"
FT FT 34..158
FT /label= D5_domain
XX
PN WO200214369-A2.
XX
XX 21-FEB-2002.
XX
XX 24-JUL-2001; 2001WO-US023185.
XX
XX 24-JUL-2000; 2000US-0220194P.
XX
XX (ANTE-) ATTENUON LLC.
XX
XX Mazar AP, Juarez JC;
XX
XX WPI; 2002-393611/42.
XX
XX Novel human kininogen D5 domain polypeptides useful for treating
XX conditions associated with endothelial cell migration, proliferation,
XX invasion or angiogenesis, e.g. arthritis, macular degeneration, benign
XX hyperplasia.
XX
XX Example 3; Page 62; 84pp; English.
XX
XX The present invention describes an isolated polypeptide (I) that
XX corresponds to the D5 domain of human kininogen, or biologically active
XX peptide fragment, homologue or functional derivative, and which: (a)
XX inhibits angiogenesis; (b) binds to the D5 binding site on endothelial
XX cells (EC); (c) activates signalling pathways leading to the introduction
XX of apoptosis in EC; and/or (d) inhibits the signalling pathway required
XX for maintenance of EC viability. (I) has cytostatic, antitumor,
XX antithrombotic, vasotrophic, vulnar, tranquiliser, thrombolytic,
XX ophthalmological, gynaecological, antiulcer, antidiabetic, antiarthritic,
XX antiangiogenic, antipapillary and endocrine activities. An antibody (IX)
XX specific for an epitope of (I) is useful for inhibiting tumour growth or
XX angiogenesis in a subject. (I), a D5 fusion polypeptide (II) or a dimeric
XX or trimeric fusion polypeptide (III) can be used for inhibiting EC
XX migration, proliferation, invasion, or angiogenesis, or for inducing EC
XX apoptosis. An angiogenic EC-targeting pharmaceutical composition (X)
XX comprising (I), (II), or (III), can be used for treating a subject having
XX a disease or condition associated with undesired EC migration,
XX proliferation, invasion or angiogenesis. (I), (II), or (III) can be used
XX for isolating a D5 domain binding molecule from a complex mixture and for
XX isolating or enriching cells expressing D5 domain binding sites from a
XX cell mixture. The present sequence represents a calmodulin binding
XX protein and D5 domain fusion protein, which is given in an example from
XX the present invention
XX
XX Sequence 158 AA;

```

```

Query Match 98.0%; Score 96; DB 5; Length 158;
Best Local Similarity 94.1%; Pred. No. 1.3e-06;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 HGHEQHQHGLGHGKFKL 17
   |||||
DB 95 HGHEQHQHGLGHGKFKL 111

RESULT 10
AAY93349
ID AAY93349 standard; peptide; 186 AA.
XX
AC AAY93349;
XX
DT 04-SEP-2000 (first entry)
XX
DE Light chain of human high molecular weight kininogen analogue.
XX
KW Human; high molecular weight kininogen; glycoprotein; endothelial cell;
KW plasma kallikrein; heavy chain; light chain; analogue; angiogenesis;
KW endothelial cell proliferation; endothelial cell migration; vitronectin.
XX
OS Synthetic.
OS Homo sapiens.
XX
XX WO200027415-A2.
XX
XX 18-MAY-2000.
XX
XX 09-NOV-1999; 99WO-US026377.
XX
XX 10-NOV-1998; 98US-0107844P.
XX
XX (UTEM) UNIV TEMPLE.
XX (DUPO) DUPONT PHARM CO.
XX (COLM/) COLMAN W R.
XX (MOUS/) MOUSA A S.
XX
XX Colman WR, Mousa AS;
XX
XX WPI; 2000-376306/32.
XX
XX Method for inhibiting endothelial cell proliferation, using compound that
XX inhibit endothelial cell migration.
XX
XX Claim 9; Page 38; 41pp; English.
XX
XX The present sequence represents an analogue of the light chain of human
XX high molecular weight kininogen. High molecular weight kininogen is a 120
XX kDa glycoprotein which binds with high affinity to endothelial cells,
XX where it is cleaved by plasma kallikrein into heavy and light chains.
XX Analogues of high molecular weight kininogen are used in the method of
XX the invention. The specification describes a method of inhibiting
XX endothelial cell proliferation. The method comprises contacting
XX endothelial cells with a compound containing high molecular weight
XX kininogen analogues. The method and the compounds can be used for
XX inhibiting endothelial cell proliferation. The compounds can also be used
XX for inhibiting angiogenesis. The compounds can also be used to inhibit
XX migration of endothelial cells to vitronectin
XX
XX Sequence 186 AA;

```

```

Query Match 98.0%; Score 96; DB 3; Length 186;
Best Local Similarity 94.1%; Pred. No. 1.5e-06;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 HGHEQHQHGLGHGKFKL 17
   |||||
DB 5 HGHEQHQHGLGHGKFKL 21

```



```

RESULT 11
AAY93342
XX AAY93342 standard; protein; 255 AA.
AC AAY93342;
XX
DT 04-SEP-2000 (first entry)
XX
DE Light chain of human high molecular weight kininogen.
XX
KW Human; high molecular weight kininogen; glycoprotein; endothelial cell;
KW Plasma kallikrein; heavy chain; light chain; analogue; angiogenesis;
KW endothelial cell proliferation; endothelial cell migration; vitronectin.
XX
OS Homo sapiens.
XX
PN WO200027415-A2.
XX
PD 18-MAY-2000.
XX
PF 09-NOV-1999; 99WO-US026377.
XX
PR 10-NOV-1998; 98US-0107844P.
XX
PA (UTEM ) UNIV TEMPLE.
PA (DUPO ) DUPONT PHARM CO.
PA (COLM/) COLMAN W R.
PA (MOUS/) MOUSA A S.
XX
PI Colman WR, Mousa AS;
XX
DR WPI; 2000-376306/32.
XX
PT Method for inhibiting endothelial cell proliferation, using compound that
PT inhibit endothelial cell migration.
XX
PS Disclosure; Page 3; 41pp; English.
XX
CC The present sequence represents the light chain of human high molecular
CC weight kininogen. High molecular weight kininogen is a 120 kDa
CC glycoprotein which binds with high affinity to endothelial cells, where
CC it is cleaved by plasma kallikrein into heavy and light chains. Analogues
CC of high molecular weight kininogen are used in the method of the
CC invention. The specification describes a method of inhibiting endothelial
CC cell proliferation. The method comprises contacting endothelial cells
CC with a compound containing high molecular weight kininogen analogues. The
CC method and the compounds can be used for inhibiting endothelial cell
CC proliferation. The compounds can also be used for inhibiting
CC angiogenesis. The compounds can also be used to inhibit migration of
CC endothelial cells to vitronectin
XX
SQ Sequence 255 AA;
XX
Query Match 98.0%; Score 96; DB 3; Length 255;
Best Local Similarity 94.1%; Pred. No. 2.1e-06;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 HGHEQQHGLGHGXKL 17
DB 74 HGHEQQHGLGHGXFKL 90
XX
RESULT 12
AAB73620
XX AAB73620 standard; protein; 255 AA.
XX
AC AAB73620;
XX
DT 10-AUG-2001 (first entry)
XX
DE Human high molecular weight kininogen (HK) light chain.
XX
KW Human; high molecular weight kininogen; HK; light chain; domain 5;

```

```

KW antibody; monoclonal antibody C11C1; neovascularisation inhibition;
KW endothelial cell proliferation inhibition; antiangiogenic;
KW vascular tube formation inhibition; diabetic retinopathy;
KW rheumatoid arthritis; atherosclerotic plaque rupture; cancer; tumour;
KW cytostatic; antidiabetic; ophthalmological; antirheumatic; antiarthritic;
KW antiatherosclerotic.
XX
OS Homo sapiens.
XX
PH Key
XX Location/Qualifiers
XX 1..12
XX Region /note= "C-terminal portion of HK domain 4 remaining after
XX bradykinin liberation"
XX Domain 13..131
XX /label= Domain 5
XX /note= "Corresponds to residues 384-502 of HK"
XX Region 69..131
XX /note= "Referred to in Claim 3; corresponds to HK
XX residues 440-502"
XX Region 69..84
XX /note= "Referred to in Claim 4; corresponds to HK
XX residues 440-457"
XX Domain 70..86
XX /label= His/Gly-rich_subdomain
XX Domain 100..125
XX /label= Endothelial_cell_binding_domain
XX Domain 104..131
XX /label= His/Gly/Lys-rich_subdomain
XX Region 115..131
XX /note= "Referred to in Claim 5; corresponds to HK
XX residues 486-502"
XX Domain 132..255
XX /label= Domain 6
XX /note= "Corresponds to residues 503-626 of HK"
XX
WO200134195-A1.
XX
PD 17-MAY-2001.
XX
XX 10-NOV-2000; 2000WO-US030975.
XX
PR 12-NOV-1999; 99US-0165165P.
XX
XX (UTEM ) UNIV TEMPLE.
XX (DUPO ) DUPONT PHARM CO.
XX
PI Colman RW, Mousa SA;
XX
XX WPI; 2001-328940/34.
XX
XX Inhibiting angiogenesis in a mammal using an antibody against high
XX molecular weight kininogen domain 5.
XX
XX Claim 4; Page 2; 38pp; English.
XX
XX The invention relates to a method of inhibiting angiogenesis in a mammal,
XX comprising administering an antibody against an epitope of high molecular
XX weight kininogen (HK) domain 5. In particular, the antibody used can be
XX monoclonal antibody C11C1 which is produced by hybridoma ATCC HB-8964.
XX The method of the invention is used to inhibit endothelial cell
XX proliferation, vascular tube formation and/or neovascularisation in
XX disease states such as diabetic retinopathy, rheumatoid arthritis and
XX atherosclerotic plaques. The antibody may be administered to prevent
XX plaque rupture, which leads to thrombotic occlusion of coronary or
XX cerebral arteries. The antibody may also be used to inhibit tumour growth
XX via the inhibition of angiogenesis. The present sequence represents human
XX high molecular weight kininogen (HK) light chain, which contains HK
XX domain 5
XX
SQ Sequence 255 AA;
XX
Query Match 98.0%; Score 96; DB 4; Length 255;
Best Local Similarity 94.1%; Pred. No. 2.1e-06;

```

Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 HGHEQOQHGLGHGHEFKL 17
 124 HGHEQOQHGLGHGHEFKL 140

Db

RESULT 14
 ABR41202
 ID ABR41202 standard; protein; 357 AA.
 XX
 AC ABR41202;
 XX
 DT 02-JUN-2003 (first entry)
 XX
 DE Human DTHP extracellular signalling protein.
 XX
 DE Human; dithp, diagnostic and therapeutic polynucleotide; diagnosis;
 KW cancer; cell proliferative disorder; autoimmune disorder;
 KW inflammatory disorder; infection; hormonal disorder; metabolic disorder;
 KW neurological disorder; gastrointestinal disorder; transport disorder;
 KW connective tissue disorder; drug screening; proteome analysis;
 KW gene therapy; antisense therapy; genotyping; transgenic animal; knock in;
 KW disease model; toxicological testing; transcript imaging;
 KW extracellular signalling.
 XX
 OS Homo sapiens.
 XX
 OS WO200297031-A2.
 XX
 PN 05-DEC-2002.
 XX
 PF 27-MAR-2002; 2002WO-US010056.
 XX
 PR 28-MAR-2001; 2001US-0279619P.
 PR 29-MAR-2001; 2001US-0280067P.
 PR 23-MAR-2001; 2001US-0280068P.
 PR 16-MAY-2001; 2001US-0291280P.
 PR 17-MAY-2001; 2001US-0291829P.
 PR 17-MAY-2001; 2001US-0291849P.
 PR 19-JUN-2001; 2001US-0299428P.
 PR 20-JUN-2001; 2001US-0299776P.
 PR 20-JUN-2001; 2001US-0300001P.
 XX
 PA (INCY-) INCYTE GENOMICS INC.
 XX
 XX Dafo A, Jones AL, Tran AB, Dahl CR, Gietzen D, Chinn J;
 PI Dufour GE, Hillman JU, Yu JY, Tuason O, Yap PE, Amshay SR;
 PI Daugherty SC, Dam TC, Liu TF, Nguyen DA, Kleefeld Y, Gerstin EH;
 PI Peralta CH, David MH, Lewis SA, Chen AT, Panzer SR, Harris B;
 PI Flores V, Marwaha R, Lo A, Lan XY, Urashka ME;
 XX
 DR WPI; 2003-129518/12.
 DR N-PSDB; ACC46146.
 XX
 XX Novel human diagnostic and therapeutic polypeptide useful for identifying
 PT test compound which specifically binds to a polypeptide encoded by human
 PT diagnostic and therapeutic polynucleotide, and to induce antibodies.
 XX
 PS Claim 27; SEQ ID NO 737; 591pp; English.
 XX
 XX The invention relates to novel human diagnostic and therapeutic
 CC polynucleotides designated dithp (ACC46080-ACC46749) and to their encoded
 CC proteins (DITHP; ABR41136-ABR41812). The invention also relates to
 CC polynucleotide sequences at least 90% identical to the dithp cDNA
 CC sequences of the invention; recombinant vectors, host cells and
 CC transgenic organisms comprising a dithp nucleic acid sequence; the
 CC recombinant production of DITHP proteins; antibodies specific for DITHP
 CC proteins; microarrays comprising dithp nucleic acid sequences; methods of
 CC detecting dithp nucleotide and protein sequences; methods of screening
 CC for compounds which specifically bind a DITHP protein; and methods of
 CC assessing the toxicity of test compounds using a dithp hybridisation
 CC probe. Dithp nucleic acid sequences and DITHP proteins may be used in the
 CC diagnosis of a wide variety of conditions including cancer and other cell
 CC proliferative disorders; autoimmune or inflammatory disorders; bacterial,

Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 HGHEQOQHGLGHGHEFKL 17
 124 HGHEQOQHGLGHGHEFKL 140

Db

RESULT 13
 ABR21100
 ID ABR21100 standard; protein; 305 AA.
 XX
 AC ABR21100;
 XX
 DT 18-FEB-2002 (first entry)
 XX
 DE Novel human diagnostic protein #21091.
 XX
 DE Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 KW
 XX Homo sapiens.
 XX
 OS WO200175067-A2.
 XX
 PN 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-US008631.
 XX
 PR 31-MAR-2000; 2000US-00540217.
 PR 23-AUG-2000; 2000US-00649167.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 XX Drmanac RT, Liu C, Tang YT;
 PI
 DR WPI; 2001-639362/73.
 DR N-PSDB; AAS85287.
 XX
 XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.
 XX
 PS Claim 20; SEQ ID NO 51459; 103pp; English.
 XX
 XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC Genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABR00010-ABG30377 represent novel human diagnostic
 CC amino acid sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 XX Sequence 305 AA;

Query Match 98.0%; Score 96; DB 4; Length 305;
 Best Local Similarity 94.1%; Pred. No. 2.5e-06;
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

CC viral, fungal or parasitic infections; hormonal disorders; metabolic
 CC disorders; neurological disorders; gastrointestinal disorders; transport
 CC disorders; and connective tissue disorders. They may also be used to
 CC screen for modulators of protein activity or gene expression. DITHP
 CC proteins can additionally be used in analysis of the proteome of a tissue
 CC or cell type and to induce antibodies. The dithp nucleic acids are
 CC additionally useful in somatic or germline gene therapy of the disorders
 CC mentioned above, as a source of antisense sequences, as a source of
 CC probes and primers, in genotyping and identification of individuals, in
 CC the generation of transgenic animal models of human disease or knock in
 CC humanised animals, in toxicological testing, and in transcript imaging.
 CC The present sequence represents a DITHP protein which has extracellular
 CC signalling activity. Note: The sequence data for this patent did not form
 CC part of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 XX
 SQ Sequence 357 AA;

Query Match 98.0%; Score 96; DB 6; Length 357;
 Best Local Similarity 94.1%; Pred. No. 3e-06;
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 HGHEQQHGLGHGKXKL 17
 |||||
 Db 176 HGHEQQHGLGHGKFKL 192

RESULT 15
 ABU99146
 ID ABU99146 standard; protein; 415 AA.

AC ABU99146;

XX
 XX
 DT 01-AUG-2003 (first entry)

DE Novel human GPCR related protein NOV12d.

XX Human; G-protein coupled receptor related protein; GPCR related protein;
 KW NOV; cytostatic; cardiant; antiarteriosclerotic; antidiabetic;
 KW immunomodulator; anti-HIV; anorectic; antiasthmatic; haemostatic;
 KW antiparkinsonian; neuroprotective; nootropic; gene therapy; vaccine;
 KW NOVX-associated disorder; cardiomyopathy; atherosclerosis; cancer;
 KW diabetes; immune disorder; AIDS; obesity; asthma;
 KW haematopoietic disorder; Parkinson's disease; Alzheimer's disease;
 KW infection; multiple sclerosis; cancer-associated cachexia;
 KW wasting disorder; chronic disease; neurogenesis; cell differentiation;
 KW cell proliferation; haematopoiesis; wound healing; angiogenesis;
 KW chromosome mapping; tissue typing; preventive medicine; pharmacogenomic.

XX Homo sapiens.

XX WO200299116-A2.

XX 12-DEC-2002.

XX 04-JUN-2002; 2002WO-US017428.

XX 04-JUN-2001; 2001US-0295607P.
 PR 04-JUN-2001; 2001US-0295661P.
 PR 06-JUN-2001; 2001US-0296404P.
 PR 06-JUN-2001; 2001US-0296418P.
 PR 14-JUN-2001; 2001US-0298285P.
 PR 15-JUN-2001; 2001US-0298556P.
 PR 21-JUN-2001; 2001US-0299492P.
 PR 26-JUN-2001; 2001US-0300883P.
 PR 28-JUN-2001; 2001US-0301550P.
 PR 13-AUG-2001; 2001US-0311972P.
 PR 27-AUG-2001; 2001US-0315071P.
 PR 29-AUG-2001; 2001US-0315660P.
 PR 14-SEP-2001; 2001US-0322293P.
 PR 17-SEP-2001; 2001US-0322706P.
 PR 14-DEC-2001; 2001US-0341186P.
 PR 28-FEB-2002; 2002US-0361189P.

PR 12-MAR-2002; 2002US-0363673P.
 PR 12-MAR-2002; 2002US-0363676P.
 PR 03-JUN-2002; 2002US-00363676.

XX (CURA-) CURAGEN CORP.

PI Anderson DW, Baumgartner JC, Boldog FL, Casman SJ, Edinger SR,
 PI Gangolli EA, Gerlach VL, Gorman L, Guo X, Hjalit T, Kekuda R, Li L,
 PI MacDougall JR, Malyankar UM, Millet I, Padigaru M, Patturajan M,
 PI Pena CE, Restelli L, Shimkets RA, Stone DJ, Spytek KA, Vernet CAM,
 PI Voss EZ, Zernusen BD;

XX WPI; 2003-140627/13.
 DR N-PSDB; ACD03650.

XX New NOVX polypeptides and nucleic acids, useful for preventing or
 PT treating NOVX-associated disorders, e.g. cancer, cardiomyopathy,
 PT atherosclerosis, or diabetes, and in chromosome mapping, tissue typing or
 PT pharmacogenomics.

XX Claim 1; Page 145; 332pp; English.

XX The invention describes an isolated polypeptide (I) comprising any of 27
 CC 118-961 residue amino acid sequences, given in the specification, a
 CC mature form of them, a sequence that is at least 95 % identical to them,
 CC or a sequence having one or more conservative substitutions in them. The
 CC polypeptide is useful in manufacturing a medicament for treating a
 CC syndrome associated with a human disease selected from a pathology
 CC associated with the polypeptide. The NOVX polypeptides, polynucleotides
 CC and antibodies are useful in treating or preventing NOVX-associated
 CC disorders, e.g. cardiomyopathy, atherosclerosis, cancer, diabetes, immune
 CC disorders, AIDS, obesity, asthma, haematopoietic disorders, Parkinson's
 CC disease, Alzheimer's disease, infections, multiple sclerosis, cancer-
 CC associated cachexia, and other wasting disorders associated with chronic
 CC diseases. The nucleic acids and polypeptides may also be used as targets
 CC for the identification of small molecules that modulate or inhibit e.g.
 CC neurogenesis, cell differentiation, cell proliferation, haematopoiesis,
 CC wound healing and angiogenesis, in gene therapy, in generation of
 CC antibodies that bind immunospecifically to NOVX substances for use in
 CC therapeutic or diagnostic methods. The nucleic acids are further used as
 CC hybridisation probes, in chromosome mapping, tissue typing, preventive
 CC medicine, and pharmacogenomics. The polypeptides are also useful as
 CC vaccines. This is the amino acid sequence of a novel human G-protein
 CC coupled receptor related protein NOV

XX Sequence 415 AA;

Query Match 98.0%; Score 96; DB 6; Length 415;
 Best Local Similarity 94.1%; Pred. No. 3.5e-06;
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 HGHEQQHGLGHGKXKL 17
 |||||
 Db 234 HGHEQQHGLGHGKFKL 250

Search completed: March 4, 2004, 13:07:26
 Job time : 49.25 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2004, 13:02:33 ; Search time 10.5 seconds
(without alignments)
155.739 Million cell updates/sec

Title: SEQ-A
Perfect score: 98
Sequence: 1 HGHQHQHGLGHGKXK1 17

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 9619:526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78: *
1: PIR1: *
2: PIR2: *
3: PIR3: *
4: PIR4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	96	98.0	644	1 KGHUHL	kininogen, HMW pre
2	81	82.7	619	1 KGBOH2	kininogen, HMW II
3	81	82.7	621	1 KGBOH1	kininogen, HMW I p
4	79	80.6	264	2 C25486	K-kininogen, HMW p
5	79	80.6	639	2 A25486	kininogen, HMW I p
6	76	77.6	290	2 C27115	K-kininogen, LMW p
7	76	77.6	315	2 A27115	major acute phase
8	62	63.3	314	2 T35241	hypothetical prote
9	62	63.3	361	2 F87286	cation efflux fan
10	59	60.2	398	2 T02681	probable zinc tran
11	58	59.2	199	2 T48099	hypothetical prote
12	58	59.2	535	2 S66148	gene pipsqueak pro
13	58	59.2	1085	2 S66149	gene pipsqueak pro
14	57	58.2	670	2 P36791	hypothetical prote
15	57	58.2	735	2 T45059	hypothetical prote
16	57	58.2	2038	2 A43742	female sterile hom
17	56	57.1	439	2 S58327	cobalt accumulat
18	55	56.1	189	2 C81428	peptidyl-prolyl ci
19	55	56.1	457	2 S39079	puif C-8 protein -
20	54	55.1	335	2 AC2686	cation efflux syst
21	54	55.1	338	2 H97467	probable inner mem
22	54	55.1	503	2 S54302	zinc transporter Z
23	54	55.1	507	2 S54303	zinc transport pro
24	53	54.1	18	2 B32473	histidine-rich pro
25	53	54.1	85	2 A45969	hemolymph antifung
26	53	54.1	110	2 T07618	cold stress protei
27	53	54.1	529	2 T08684	hypothetical prote
28	52	53.1	378	2 T49164	zinc transporter-1
29	52	53.1	389	2 B96635	hypothetical prote

ALIGNMENTS

RESULT 1

KGHUHL

kininogen, HMW precursor [validated] - human
N;Alternate names: alpha-2-thiol proteinase inhibitor; preprokininogen, prokininogen
C;Species: Homo sapiens (man)
C;Date: 28-May-1986 #sequence revision 28-May-1986 #text change 08-Dec-2000
C;Accession: A01279; A25276; S32422; A91153; A24871; A27899; A27699; A31905; A34030; S024
R;Okubo, I.; Kurachi, K.; Takasawa, T.; Shiohara, H.; Sasaki, M.
Biochemistry 23, 5691-5697, 1984
A;Title: Isolation of a human cDNA for alpha-2-thiol proteinase inhibitor and its identit
A;Reference number: A90490; MUID:85122621; PMID:6441591
A;Accession: A01279
A;Molecule type: mRNA
A;Residues: 1-389 <CHK>
A;Cross-references: GB:K02566; NID:g177889
R;Takagaki, Y.; Kitamura, N.; Nakanishi, S.
J. Biol. Chem. 260, 8601-8609, 1985
A;Title: Cloning and sequence analysis of cDNAs for human high molecular weight and low m
A;Reference number: A92544; MUID:85234582; PMID:2989293
A;Accession: A25276
A;Molecule type: mRNA
A;Residues: 1-592, 'I', 594-644 <TAX>
A;Cross-references: GB:M11437; NID:g186751; PIDN:AAB59550.1; PID:g386852
R;Auerwald, E.A.; Roessler, D.; Mentele, R.; Assfalg-Machleidt, I.
FEBS Lett. 321, 93-97, 1993
A;Title: Cloning, expression and characterization of human kininogen domain 3.
A;Reference number: S32422; MUID:93223854; PMID:8467916
A;Accession: S32422
A;Molecule type: mRNA
A;Residues: 'ANSM', 253-377 <AUE>
A;Note: differences are due to known cloning artifacts
R;Lottspeich, F.; Kellermann, J.; Henschen, A.; Foertsch, B.; Muller-Esterl, W.
Eur. J. Biochem. 152, 307-311, 1985
A;Title: The amino acid sequence of the light chain of human high-molecular-mass kininoge
A;Reference number: A91153; MUID:86030270; PMID:4054110
A;Accession: A91153
A;Molecule type: protein
A;Residues: 379-644 <LOT>
A;Note: the bradykinin sequence preceding the light chain sequence was not determined in
R;Kellermann, J.; Lottspeich, F.; Henschen, A.; Mueller-Esterl, W.
Eur. J. Biochem. 154, 471-478, 1986
A;Title: Completion of the primary structure of human high-molecular-mass kininogen. The
A;Reference number: A24871; MUID:86108361; PMID:3484703
A;Accession: A24871
A;Molecule type: protein
A;Residues: 'Z', 20-380 <KEIL>
R;Kellermann, J.; Lottspeich, F.; Henschen, A.; Mueller-Esterl, W.
in Kinins IV, Greenbaum, L.M., and Margolius, H.S., ed., pp.85-89, Plenum Press, New York
A;Title: Amino acid sequence of the light chain of human high molecular mass kininogen.
A;Reference number: A27899
A;Accession: A27899

hypothetical prote
conserved hypothet
hypothetical prote
hypothetical prote
ATP/GTP-binding pr
nuclear distributi
NADH dehydrogenase
hypothetical prote
chorion protein s3
hypothetical prote
hypothetical prote
hypothetical prote
zinc transporter-1
urease accessory p
hypothetical prote

30 51.5 52.6 410 2 T26757
31 51 52.0 106 2 E87560
32 51 52.0 160 2 T07180
33 51 52.0 390 2 T34137
34 51 52.0 409 2 E83992
35 51 52.0 697 2 T03834
36 50.5 51.5 686 2 F87489
37 50 51.0 549 2 T15506
38 49.5 50.5 306 2 S08607
39 49 50.0 66 2 B83515
40 49 50.0 203 2 T36240
41 48 49.0 73 2 H95898
42 48 49.0 179 2 A85217
43 48 49.0 334 2 T47986
44 47.5 48.5 171 2 AE3458
45 47 48.0 224 2 T34937

A:Molecule type: protein
 A:Residues: 379-389, 'K', 390-407, 'Q', 409-644 <KEL2>
 R:Mindroiu, T.; Carretero, O.A.; Proulx, D.; Walz, D.; Scicli, A.G.
 Biochem. Biophys. Res. Commun. 152, 519-526, 1988
 A:Title: A new kinin moiety in human plasma kininogens.
 A:Reference number: A27699; MUID:88209021; PMID:3365237
 A:Accession: A27699
 A:Molecule type: protein
 A:Residues: 380-389 <MIN>
 R:Maeda, H.; Matsumura, Y.; Kato, H.
 J. Biol. Chem. 263, 16051-16054, 1988
 A:Title: Purification and identification of [hydroxyprolyl(3)]bradykinin in ascitic fluid
 A:Reference number: A31905; MUID:89034061; PMID:3182782
 A:Accession: A31905
 A:Molecule type: protein
 A:Residues: 381-389 <MAB>
 R:Sasaguri, M.; Ikeda, M.; Ideishi, M.; Arakawa, K.
 Biochem. Biophys. Res. Commun. 150, 511-516, 1988
 A:Title: Identification of [hydroxyproline(3)]-lysyl-bradykinin released from human plas
 A:Reference number: A34030; MUID:88106632; PMID:3337729
 A:Accession: A34030
 A:Molecule type: protein
 A:Residues: 380-389 <SAS>
 R:Lenarcic, B.; Gabrijelcic, D.; Rozman, B.; Drobnic-Kosorok, M.; Turk, V.
 Biol. Chem. Hoppe-Seyler 369, 257-261, 1988
 A:Title: Human cathepsin B and cysteine proteinase inhibitors (CPIs) in inflammatory and
 A:Reference number: S02482; MUID:89076517; PMID:3264507
 A:Accession: S02482
 A:Molecule type: protein
 A:Residues: 1-19; 189-192; 310-314; 381-389 <LEN1>
 R:Kato, H.; Matsumura, Y.; Maeda, H.
 FEBS Lett. 222, 252-254, 1988
 A:Title: Isolation and identification of hydroxyproline analogues of bradykinin in human
 A:Reference number: A61495; MUID:88211869; PMID:3366244
 A:Accession: A61495
 A:Molecule type: protein
 A:Residues: 380-389 <KAT1>
 A:Experimental source: urine
 A:Accession: B61495
 A:Molecule type: protein
 A:Residues: 381-389 <KAT2>
 A:Experimental source: urine
 A:Note: this peptide had Pro-383 modified to 4-hydroxyproline
 A:Accession: C61495
 A:Molecule type: protein
 A:Residues: 380-389 <KAT3>
 R:Lenarcic, B.; Krasovec, M.; Ritonja, A.; Olafsson, I.; Turk, V.
 FEBS Lett. 280, 211-215, 1991
 A:Title: Inactivation of human cystatin C and kininogen by human cathepsin D.
 A:Reference number: S14303; MUID:91192133; PMID:2013314
 A:Accession: S14447
 A:Molecule type: protein
 A:Residues: 264-359, 'N', 361-375 <LEN2>
 R:Little, S.S.; Johnson, D.A.
 Biochem. J. 307, 341-345, 1995
 A:Title: Human mast cell tryptase isoforms: separation and examination of substrate-spec
 A:Reference number: S55339; MUID:95251593; PMID:7733867
 A:Accession: S55339
 A:Molecule type: protein
 A:Residues: 450-452, 'X', 454, 'X', 456 <LIT>
 R:Straczek, J.; Maachi, F.; le Nguyen, D.; Becchi, M.; Heulin, M.H.; Nabet, P.; Bellevil
 FEBS Lett. 373, 207-211, 1995
 A:Title: Purification from human plasma of a tetrapeptide that potentiates insulin-like
 A:Reference number: S68059; MUID:96033974; PMID:7589467
 A:Accession: S68059
 A:Molecule type: protein
 A:Residues: 431-434 <STR>
 R:Kitamura, N.; Kitagawa, H.; Fukushima, D.; Takagaki, Y.; Miyata, T.; Nakanishi, S.
 J. Biol. Chem. 260, 8610-8617, 1985
 A:Title: Structural organization of the human kininogen gene and a model for its evoluti
 A:Reference number: A92545; MUID:85234583; PMID:2989294
 A:Contents: annotation; gene organization

R:Pierce, J.V.
 Fed. Proc. 27, 52-57, 1968
 A:Title: Structural features of plasma kinins and kininogens.
 A:Reference number: A91455; MUID:90255622; PMID:4952632
 A:Contents: annotation; bradykinin
 C:Comment: The HMW kininogen precursor and the LMW form are produced from the same gene t
 C:Comment: Kininogen is a cysteine proteinase inhibitor, takes part in initiation of the
 C:Comment: The glycine/histidine/lysine-rich region of HMW kininogen light chain is impor
 C:Comment: Bradykinin, released from kininogen by kallikrein, is a potent vasodilator, ir
 xyproline residue is present in the kininogen prior to the release of bradykinin.
 A:Genetics:
 A:Gene: GDB:KMG
 A:Cross-references: GDB:125256; OMIM:228960
 A:Map position: 3q27-3q27
 A:Introns: 65/3; 102/3; 131/1; 188/3; 224/3; 253/1; 310/3; 346/3; 375/3
 A:Superfamily: kininogen; cystatin homology
 C:Keywords: alternative splicing; blood coagulation; cysteine proteinase inhibitor; dupli
 F:1-18/Domain: signal sequence #status experimental <SIG>
 F:19-644/Product: HMW kininogen I (prokininogen) #status experimental <MAT1>
 F:19-379/Domain: HMW kininogen II #status experimental <MAT2>
 F:19-379/Domain: HMW kininogen heavy chain #status experimental <HCH>
 F:19-131/Domain: cystatin homology <CY1>
 F:142-253/Domain: cystatin homology <CY2>
 F:264-375/Domain: cystatin homology <CY3>
 F:380-389/Product: lysyl-bradykinin (kallidin II) #status experimental <KBDY>
 F:381-389/Product: bradykinin (kallidin I) #status experimental <BDY>
 F:390-644/Product: HMW kininogen light chain #status experimental <LCH>
 F:421-510/Region: glycine/histidine/lysine-rich 30-residue repeats
 F:431-434/Product: low molecular weight growth promoting factor #status experimental <GPI
 F:19/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experiment
 F:28-614, 83-94, 107-126, 142-145, 206-218, 229-248, 264-267, 328-340, 351-370/bisulfide bonds: #
 F:48/Binding site: carboxylate (Asn) (covalent) #status absent
 F:169, 205, 294/Binding site: carboxylate (Asn) (covalent) #status experimental
 F:379-380/Cleavage site: Met-Lys (kallikrein) #status experimental
 F:383/Modified site: 4-hydroxyproline (Pro) (partial) #status experimental
 F:389-390/Cleavage site: Arg-Ser (kallikrein) #status experimental
 F:401, 533, 542, 546, 557, 571, 593, 628/Binding site: carboxylate (Thr) (covalent) #status ex
 F:577/Binding site: carboxylate (Ser) (covalent) #status experimental
 Query Match 98.0%; Score 96; DB 1; Length 644;
 Best Local Similarity 94.1%; Pred. No. 7.9e-07;
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 HGHEQHQHGLGHGHKXKL 17
 DB 463 HGHEQHQHGLGHGHKFKL 479
 RESULT 2
 KGBORZ
 kininogen, HMW II precursor - bovine
 N:Alternate names: alpha-2-thiol proteinase inhibitor; preprokininogen
 N:Contains: bradykinin (kallidin); kininogen I; kininogen II; prokininogen
 C:Species: Bos primigenius taurus (cattle)
 C>Date: 14-Nov-1983 #sequence revision 14-Nov-1983 #text_change 22-Jun-1999
 C:Accession: A01282; A91923; A91941; A91938; B29559
 R:Kitamura, N.; Takagaki, Y.; Furuto, S.; Tanaka, T.; Nawa, H.; Nakanishi, S.
 Nature 305, 545-549, 1983
 A:Title: A single gene for bovine high molecular weight and low molecular weight kininoge
 A:Reference number: A93317; MUID:84014106; PMID:6571699
 A:Accession: A01282
 A:Molecule type: mRNA
 A:Residues: 1-619 <KIT>
 A:Cross-references: GB:V01492; GB:K01758; NID:g493; PIDN:CAA24736.1; PID:g494
 R:Kato, H.; Nagasawa, S.; Suzuki, T.
 J. Biochem. 67, 313-323, 1970
 A:Title: Studies on the structure of bovine kininogen: cleavages of disulfide bonds and c
 A:Reference number: A91923; MUID:70180420; PMID:4986212
 A:Accession: A91923
 A:Molecule type: protein
 A:Residues: 376-391 <KAT>
 R:Han, Y.N.; Kato, H.; Iwanaga, S.; Suzuki, T.
 J. Biochem. 79, 1201-1222, 1976

N:Alternate names: alpha-2-thiol proteinase inhibitor; preprokinogen
N:Contains: bradykinin (kallidin); kininogen I; kininogen II; prokininogen
C:Species: Bos primigenius taurus (cattle)
C>Date: 14-Nov-1983 #sequence revision 14-Nov-1983 #text_change 22-Jun-1999
C:Accession: A01281; A91283; A91938; A29559
R:Kitamura, N.; Takaaki, Y.; Furuto, S.; Tanaka, T.; Nawa, H.; Nakanishi, S.
Nature 305, 545-549, 1983
A:Title: A single gene for bovine high molecular weight and low molecular weight kininogen
A:Reference number: A93317; MUID:84014106; PMID:6571699
A:Accession: A01281
A:Molecule type: mRNA
A:Residues: 1-621 <KT>
A:Cross-references: GB:V01491; GB:K01757; MID:g4931; PIDN:CAA24735.1; PID:g492
R:Kato, H.; Nagasawa, S.; Suzuki, T.
J. Biochem. 67 313-323, 1970
A:Title: Studies on the structure of bovine kininogen: cleavages of disulfide bonds and
A:Reference number: A91923; MUID:70180420; PMID:4986212
A:Accession: A91923
A:Molecule type: protein
A:Residues: 378-393 <KAT>
R:Han, Y.N.; Komiya, M.; Iwanaga, S.; Suzuki, T.
J. Biochem. 77, 55-68, 1975
A:Title: Studies on the primary structure of bovine high-molecular-weight kininogen. Amr
A:Reference number: A91938; MUID:75170265; PMID:1169237
A:Accession: A91938
A:Molecule type: protein
A:Residues: 458-498 <HAN>
R:Sueyoshi, T.; Miyata, T.; Hashimoto, N.; Kato, H.; Hayashida, H.; Miyata, T.; Iwanaga, S.
J. Biol. Chem. 262, 2768-2779, 1987
A:Title: Bovine high molecular weight kininogen. The amino acid sequence, positions of ca
A:Reference number: A92627; MUID:87137530; PMID:3546295
A:Accession: A29559
A:Molecule type: protein
A:Residues: 71, 20-123, 1', 125-127, 1', 129-378 <SUE>
R:Lottspeich, F.; Kellermann, J.; Henschel, A.; Foerisch, B.; Muller-Esterl, W.
Eur. J. Biochem. 152, 307-314, 1985
A:Title: The amino acid sequence of the light chain of human high-molecular-mass kininogen
A:Reference number: A91153; MUID:86030270; PMID:4054110
A:Contents: annotation; bovine cleavage sites; bovine carbohydrate binding sites
R:Sueyoshi, T.; Miyata, T.; Kato, H.; Iwanaga, S.
Sekigaku 56, 808, 1984
A:Title: Disulfide bonds in bovine HMW kininogens.
A:Reference number: A94300
A:Contents: annotation; disulfide bonds
A:Note: article in Japanese
C:Comment: The HMW kininogen precursor is produced from the same gene as the LMW form as
C:Comment: Kininogen is a cysteine proteinase inhibitor, takes part in initiation of the
C:Comment: The glycine/histidine/lysine-rich region of HMW kininogen light chain is impor
C:Comment: Bradykinin, released from kininogen by kallikrein, is a potent vasodilator, ir
xyproline residue is present in the kininogen prior to the release of bradykinin.
C:Superfamily: kininogen; cystatin homology
C:Keywords: alternative splicing; blood coagulation; cysteine proteinase inhibitor; dupl
F:19-Domain: signal sequence #status predicted <SIG>
F:19-621/Product: HMW prokininogen I #status predicted <MAT>
F:19-379/Product: HMW kininogen I heavy chain #status experimental <HCH>
F:19-130/Domain: cystatin homology <CY1>
F:141-252/Domain: cystatin homology <CV2>
F:263-374/Domain: cystatin homology <CV3>
F:379-388/Product: lysyl-bradykinin (kallidin II) #status experimental <KBDY>
F:380-388/Product: bradykinin (kallidin I) #status experimental <BDY>
F:389-621/Product: HMW kininogen I light chain #status experimental <LCH>
F:19/Modified site: glycine/histidine/lysine-rich
F:19/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experiment
F:27-591, 82-93, 106-125, 141-144, 205-217, 228-247, 263-266, 327-339, 350-369/Disulfide bonds: #
F:87, 168, 169, 204/Binding site: carboxylate (Asn) (covalent) #status experimental
F:136/Binding site: carboxylate (Thr) (covalent) (partial) #status experimental
F:197/Binding site: carboxylate (Asn) (covalent) (partial) #status experimental
F:378-379/Cleavage site: Met-Lys (kallikrein) #status experimental
F:382/Modified site: 4-hydroxyproline (Pro) #status predicted
F:388-389/Cleavage site: Arg-Ser (kallikrein) #status experimental
F:398, 406, 512/Binding site: carboxylate (Ser) (covalent) #status experimental
F:399, 400, 520, 524, 536, 548, 553, 570/Binding site: carboxylate (Thr) (covalent) #status exp
F:498-499/Cleavage site: Arg-Thr (kallikrein) #status experimental

F:389-621/Product: HMW kininogen I light chain #status experimental <LCH>
F:417-448/Region: glycine/histidine/lysine-rich
F:19/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experimental
F:27-591,82-93,106-125,141-144,205-217,228-247,263-266,327-339,350-369/Disulfide bonds: #
F:87,168,189,204/Binding site: carboxylate (Asn) (covalent) #status experimental
F:136/Binding site: carboxylate (Thr) (covalent) (partial) #status experimental
F:197/Binding site: carboxylate (Asn) (covalent) (partial) #status experimental
F:378-379/Cleavage site: Met-lys (kallikrein) #status experimental
F:382/Modified site: 4-hydroxyproline (Pro) #status predicted
F:388-389/Cleavage site: Arg-Ser (kallikrein) #status experimental
F:398,406,512/Binding site: carboxylate (Ser) (covalent) #status experimental
F:399,400,520,524,536,548,553,570/Binding site: carboxylate (Thr) (covalent) #status experimental
F:498-499/Cleavage site: Arg-Thr (kallikrein) #status experimental

Query Match 82.7%; Score 81; DB 1; Length 621;
 Best Local Similarity 85.7%; Pred. No. 0.00014;
 Matches 12; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 HGHEQOQHGLGHGK 14
 |||:|||||
 Db 463 HGHQKHGLGHGK 476
 |||:|||||

RESULT 4
 C25486
 K-kininogen, HMW precursor - rat (fragment)
 C:Species: Rattus norvegicus (Norway rat)
 C>Date: 08-Mar-1989 #sequence_revision 08-Mar-1989 #text_change 30-Sep-1993
 C:Accession: C25486
 R.Kitagawa, H.; Kitamura, N.; Hayashida, H.; Miyata, T.; Nakanishi, S.
 J. Biol. Chem. 262, 2190-2198, 1987
 A>Title: Differing expression patterns and evolution of the rat kininogen gene family.
 A:Reference number: A92625; MUID:87137443; PMID:3029068
 A:Accession: C25486
 A:Molecule type: DNA
 A:Residues: 1-264 <KIT>
 A:Comment: The nucleotide sequence was obtained from GenBank, release 55.0.
 C:Superfamily: kininogen; cystatin homology

Query Match 80.6%; Score 79; DB 2; Length 264;
 Best Local Similarity 70.6%; Pred. No. 0.00011;
 Matches 12; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 HGHEQOQHGLGHGK 17
 |||:|||||
 Db 75 HGHQKHGLGHGK 91
 |||:|||||

RESULT 5
 A25486
 kininogen, HMW I precursor - rat
 N:Contains: bradykinin
 C:Species: Rattus norvegicus (Norway rat)
 C>Date: 08-Mar-1989 #sequence_revision 08-Mar-1989 #text_change 15-Nov-1996
 C:Accession: A25486
 R.Kitagawa, H.; Kitamura, N.; Hayashida, H.; Miyata, T.; Nakanishi, S.
 J. Biol. Chem. 262, 2190-2198, 1987
 A>Title: Differing expression patterns and evolution of the rat kininogen gene family.
 A:Reference number: A92625; MUID:87137443; PMID:3029068
 A:Accession: A25486
 A:Molecule type: mRNA
 A:Residues: 1-639 <KIT>
 A>Note: the authors translated the codon CAA for residue 347 as Asn
 C:Superfamily: kininogen; cystatin homology
 C:Keywords: alternative splicing
 F:1-18/Domain: signal sequence #status predicted <SIG>
 F:19-639/Product: kininogen, HMW I #status predicted <MAT>
 F:19-131/Domain: cystatin homology <CY1>
 F:142-253/Domain: cystatin homology <CY2>
 F:264-375/Domain: cystatin homology <CY3>

Query Match 80.6%; Score 79; DB 2; Length 639;
 Best Local Similarity 70.6%; Pred. No. 0.00028;
 Matches 12; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 HGHEQOQHGLGHGK 17
 |||:|||||
 Db 450 HGHQKHGLGHGK 466
 |||:|||||

RESULT 6
 C27115
 K-kininogen, LMW precursor - rat (fragments)
 C:Species: Rattus norvegicus (Norway rat)
 C>Date: 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change 20-Aug-1999
 C:Accession: C27115; A25488

R.Fung, W.P.; Schreiber, G.
 J. Biol. Chem. 262, 9298-9308, 1987
 A>Title: Structure and expression of the genes for major acute phase alpha-1-protein (thi
 A:Reference number: A92653; MUID:87250580; PMID:2439509
 A:Accession: C27115
 A:Molecule type: DNA
 A:Residues: 1-290 <FUN>
 R.Kageyama, R.; Kitamura, N.; Ohkubo, H.; Nakanishi, S.
 J. Biol. Chem. 262, 2345-2351, 1987
 A>Title: Differing utilization of homologous transcription initiation sites of rat K and
 A:Reference number: A25488; MUID:87137465; PMID:3818598
 A:Accession: A25488
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-48 <XAG>
 A:Cross-references: GB:J02662; NID:9205071; PIDN:AAA41483.1; PID:9205072
 C:Superfamily: kininogen; cystatin homology
 F:19-65/Domain: cystatin homology (fragment) <CYS>

Query Match 77.6%; Score 76; DB 2; Length 290;
 Best Local Similarity 70.6%; Pred. No. 0.00036;
 Matches 12; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 HGHEQOQHGLGHGK 17
 |||:|||||
 Db 124 NGHQKHGLGHGK 140
 |||:|||||

RESULT 7
 A27115
 major acute phase alpha-1 protein 1 - rat (fragments)
 C:Species: Rattus norvegicus (Norway rat)
 C>Date: 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change 16-Jul-1999
 C:Accession: A27115
 R.Fung, W.P.; Schreiber, G.
 J. Biol. Chem. 262, 9298-9308, 1987
 A>Title: Structure and expression of the genes for major acute phase alpha-1-protein (thi
 A:Reference number: A92653; MUID:87250580; PMID:2439509
 A:Accession: A27115
 A:Molecule type: DNA
 A>Status: not compared with conceptual translation
 A:Residues: 1-315 <FUN>
 C:Genetics: MAP1
 C:Superfamily: kininogen; cystatin homology
 F:19-65/Domain: cystatin homology (fragment) <CYS>

Query Match 77.6%; Score 76; DB 2; Length 315;
 Best Local Similarity 70.6%; Pred. No. 0.00039;
 Matches 12; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 HGHEQOQHGLGHGK 17
 |||:|||||
 Db 148 NGHQKHGLGHGK 164
 |||:|||||

RESULT 8
 T35241
 hypothetical protein SC5C7.34 SC5C7.34 - Streptomyces coelicolor
 C:Species: Streptomyces coelicolor
 C>Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 05-Nov-1999
 C:Accession: T35241
 R.Seeger, K.J.; Harris, D.; Parkhill, J.; Barrall, B.G.; Rajandream, M.A.
 submitted to the EMBL Data Library, September 1998
 A:Reference number: Z21572
 A:Accession: T35241
 A>Status: preliminary; translated from GB/EMBL/DDBU
 A:Molecule type: DNA
 A:Residues: 1-314 <SEEL>
 A:Cross-references: EMBL:AL031515; PIDN:CAA20646.1; GSPDB:GN00070; SCOEDB:SC5C7.34
 A:Experimental source: strain A3(2)
 C:Genetics:
 A:Gene: SCOEDB:SC5C7.34

Query Match 63.3%; Score 62; DB 2; Length 314;
 Best Local Similarity 76.9%; Pred. No. 0.05;
 Matches 10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 HGHEQOQHGLGHGH 13
 DB 18 HGHGQGHGHGHGH 30

RESULT 9

F87286
 cation, efflux family protein [imported] - Caulobacter crescentus
 C:Species: Caulobacter crescentus
 C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001
 C:Accession: F87286
 R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
 B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
 D.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
 Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
 A:Title: Complete Genome Sequence of Caulobacter crescentus.
 A:Reference number: A87249; MUID:21173698; PMID:11259647
 A:Accession: F87286
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-361 <STO>
 A:Cross-references: GB:AF005673; NID:gl3421446; PIDN:AAK22290.1; GSPDB:GN00148
 C:Genetics:
 A:Gene: CC0303

Query Match 63.3%; Score 62; DB 2; Length 361;
 Best Local Similarity 69.2%; Pred. No. 0.057;
 Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 HGHEQOQHGLGHGH 13
 DB 39 HGHGHAHGHGHGH 51

RESULT 10

T02681
 probable zinc transporter At2g46800 [imported] - Arabidopsis thaliana
 N:Alternate names: hypothetical protein F19D11.8
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 16-Feb-2001
 C:Accession: T02681; D84907
 R:Rounsley, S.D.; Lin, X.; Kaul, S.; Shea, T.P.; Fujii, C.Y.; Shen, M.; Ron
 submitted to the EMBL Data Library, September 1998
 A:Description: Arabidopsis thaliana chromosome II BAC F19D11 genomic sequence.
 A:Reference number: Z14698
 A:Accession: T02681
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-398 <ROU>
 A:Cross-references: EMBL:AC005310; NID:g3510247; PID:g3510254
 A:Experimental source: cultivar Columbia
 R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
 M.; Koo, H.; Morfitt, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.
 euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
 Nature 402, 761-768, 1999
 A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
 A:Reference number: A84420; MUID:20083487; PMID:10617197
 A:Accession: D84907
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-398 <STO>
 A:Cross-references: GB:AE000203; NID:g3510254; PIDN:AAK33498.1; GSPDB:GN00139
 C:Genetics:
 A:Gene: At2g46800; F19D11.8
 A:Map position: 2

Query Match 60.2%; Score 59; DB 2; Length 398;
 Best Local Similarity 69.2%; Pred. No. 0.18;

Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 HGHEQOQHGLGHGH 13
 DB 184 HGHSHGHGHGHGH 196

RESULT 11

T48099
 hypothetical protein T20010.200 - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
 C:Accession: T48099
 R:Obermaier, B.; Ottenwaelder, B.; Duchemin, D.; Zeidler, K.; Mewes, H.W.; Rudd, S.; Lemc
 submitted to the protein Sequence Database, April 2000
 A:Reference number: Z24484
 A:Accession: T48099
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-199 <OBE>
 A:Cross-references: EMBL:AL163816
 A:Experimental source: cultivar Columbia; BAC clone T20010
 C:Genetics:
 A:Map position: 3
 A:Introns: 163/2
 A:Note: T20010.200

Query Match 59.2%; Score 58; DB 2; Length 199;
 Best Local Similarity 69.2%; Pred. No. 0.12;

Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 HGHEQOQHGLGHGH 13
 DB 92 HGHGCHGHGHGHGH 104

RESULT 12

S66148
 gene pipsqueak protein A short form - fruit fly (Drosophila melanogaster)
 C:Species: Drosophila melanogaster
 C:Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 21-Jul-2000
 C:Accession: S66148
 R:Weber, U.; Siegel, V.; Mlodzik, M.
 EMBO J. 14, 6247-6257, 1995
 A:Title: pipsqueak encodes a novel nuclear protein required downstream of seven-up for ti
 A:Reference number: S66148; MUID:96134923; PMID:8557044
 A:Accession: S66148
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-535 <WEB>
 A:Cross-references: EMBL:X90986; NID:gl149498; PIDN:CAA62473.1; PID:gl149499
 C:Genetics:
 A:Gene: pipsqueak
 A:Superfamily: POZ domain homology
 F:21-123/Domain: POZ domain homology <POZ>

Query Match 59.2%; Score 58; DB 2; Length 535;
 Best Local Similarity 69.2%; Pred. No. 0.34;

Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 HGHEQOQHGLGHGH 13
 DB 332 HBEHGHGHGHGHGH 344

RESULT 13

S66149
 gene pipsqueak protein A long form - fruit fly (Drosophila melanogaster)
 C:Species: Drosophila melanogaster
 C:Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 23-Sep-2002
 C:Accession: S66149; S66150; T45461
 R:Weber, U.; Siegel, V.; Mlodzik, M.
 EMBO J. 14, 6247-6257, 1995

A:Title: pipsqueak encodes a novel nuclear protein required downstream of seven-up for
 A:Reference number: S66149; MUID:96134923; PMID:8557044

A:Accession: S66149
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1085 <WEB>

A:Cross-references: EMBL:X90986; NID:g1149498; PIDN:CAA62474.1; PID:g1149500
 A:Accession: S66150

A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 'MQ', 428-1085 <W2>
 A:Cross-references: EMBL:X90986; NID:g1149498; PIDN:CAA62475.1; PID:g1149501

R:Horowitz, H.; Berg, C.A.
 Development 122, 1859-1871, 1996
 A:Title: The Drosophila pipsqueak gene encodes a nuclear BTB-domain-containing protein

A:Reference number: Z22972; MUID:96232300; PMID:8674425
 A:Accession: T45461

A:Status: preliminary; translated from GE/EMBL/DDBJ
 A:Molecule type: mRNA

A:Residues: 1-355, 'E', 357-1005, 'H', 1007-1020, 'Q', 1021-1061, 'ERS', <HOR>
 A:Cross-references: EMBL:U48358; NID:g1203906; PIDN:ARC47153.1; PID:g1203907

A:Experimental source: tissue type ovarian
 C:Genetics:

A:Gene: pipsqueak; psq
 A:Map position: II
 A:Introns: 427/3

A:Function:
 C:Description: required for establishing polarity of the developing egg chamber

C:Superfamily: BRCOR-Z protein; POZ domain homology
 F:21-123/Domain: POZ domain homology <POZ>

Query Match 59.2%; Score 58; DB 2; Length 1085;
 Best Local Similarity 69.2%; Pred. No. 0.7;
 Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 HGHEQOQHGLGHGH 13
 |||||
 Db 332 HEHNNHGHGHGH 344

RESULT 14
 F36791

hypothetical protein ORF50 - ictaluriid herpesvirus 1 (strain auburn 1)
 C:Species: ictaluriid herpesvirus 1

A>Note: host Ictalurus punctatus (channel catfish)
 C:Date: 17-Aug-1992 #sequence_revision 17-Aug-1992 #text_change 31-Jan-2000

C:Accession: F36791
 R:Davidson, A.J.

submitted to GenBank, January 1992
 A:Description: Channel catfish virus: a new type of herpesvirus.

A:Reference number: A36804
 A:Accession: F36791

A:Molecule type: DNA
 A:Residues: 1-670 <DAV>

A:Cross-references: GB:M75136; NID:g331209; PIDN:AAA88153.1; PID:g331260
 R:Davidson, A.J.

Virology 186, 9-14, 1992
 A:Title: Channel catfish virus: a new type of herpesvirus.

A:Reference number: A39447; MUID:92087490; PMID:1727613
 A:Contents: annotation

A>Note: neither protein nor nucleic acid sequence is given
 C:Genetics:

A:Gene: 50
 C:Superfamily: period clock protein; EGF homology

Query Match 58.2%; Score 57; DB 2; Length 670;
 Best Local Similarity 69.2%; Pred. No. 0.61;
 Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 HGHEQOQHGLGHGH 13
 |||||
 Db 637 HGHGHGHGHGHGH 649

RESULT 15

T45059

hypothetical protein Y39B6B.gg [imported] - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans

C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jul-2000
 C:Accession: T45059

R:Wilson, R.; Ainscough, R.; Anderson, K.; Baynes, C.; Berks, M.; Bonfield, J.; Burton, C.
 raser, A.; Fulton, L.; Gardner, A.; Green, P.; Hawkins, T.; Hillier, L.; Jier, M.; Johnston
 B.; O'Callaghan, M.; Parsons, J.; Percy, C.; Rifken, L.; Roopra, A.; Saunders, D.

Nature 368, 32-38, 1994
 A:Authors: Showkneen, R.; Sims, M.; Smaiden, N.; Smith, A.; Smith, M.; Sonhammer, E.; St
 tock, L.; Wilkinson-Spratt, J.; Wohlman, P.

A:Title: 2.2 Mb of contiguous nucleotide sequence from chromosome III of C. elegans.
 A:Reference number: S43531; MUID:94150718; PMID:7906398

A:Accession: T45059
 A:Status: preliminary; translated from GE/EMBL/DDBJ

A:Molecule type: DNA
 A:Residues: 1-735 <WIL>

A:Cross-references: EMBL:AL132896; NID:g6434440; PIDN:CAB60938.1; PID:g6434473
 A:Experimental source: clone Y39B6B

C:Genetics:
 A:Map position: 3
 A:Introns: 18/1; 69/1

A>Note: Y39B6B.gg
 Query Match 58.2%; Score 57; DB 2; Length 735;
 Best Local Similarity 61.5%; Pred. No. 0.67;
 Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 HGHEQOQHGLGHGH 13
 |||||
 Db 603 HGHGHGHGHGHGH 615

Search completed: March 4, 2004, 13:11:21
 Job time : 11.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2004, 12:53:12 ; Search time 6.75 Seconds
(without alignments)

131.140 Million cell updates/sec

Title: SEQ-A
Perfect score: 98
Sequence: 1 HGHEQQHGLGHGKXK1 17

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	96	98.0	644	1	KNG_HUMAN
2	81	82.7	619	1	KNG2_BOVIN
3	81	82.7	621	1	KNH1_BOVIN
4	79	80.6	639	1	KNG_RAT
5	79	80.6	661	1	KNG_MOUSE
6	64	65.3	693	1	CAUP_DROME
7	59	60.2	118	1	SI09_RABIT
8	57	58.2	670	1	VG50_HSVII
9	57	58.2	2038	1	FSH_DROME
10	56	57.1	439	1	COT1_YEAST
11	54	55.1	503	1	ZNT1_MOUSE
12	54	55.1	507	1	ZNT1_RAT
13	53	54.1	85	1	ANTF_SARPE
14	53	54.1	469	1	KE4_HUMAN
15	52	53.1	389	1	INT3_ARATH
16	51	52.0	409	1	DP42_BACHD
17	49.5	50.5	306	1	CH38_DROME
18	49	50.0	449	1	CSUP_DROME
19	48	49.0	212	1	SYD1_AERHY
20	48	49.0	352	1	KE4_BRARE
21	47	48.0	155	1	KE4_PIG
22	47	48.0	515	1	KE3L_CABEL
23	47	48.0	527	1	HSF8_LYCPE
24	47	48.0	722	1	HNW2_DROME
25	47	48.0	955	1	FRU_DROME
26	46.5	47.4	147	1	YV59_CABEL
27	46.5	47.4	476	1	KE4_MOUSE
28	46	46.9	568	1	DISC_DROME
29	46	46.9	732	1	YH3_SCHPO
30	45	45.9	302	1	HYPB_BRAJA
31	45	45.9	317	1	YG98_MYCLE
32	45	45.9	336	1	FILA_MOUSE
33	45	45.9	351	1	HRPX_PLALO

34	45	45.9	421	1	PO41_MOUSE
35	44.5	45.4	635	1	HMLA_DROME
36	44	44.9	274	1	YOEM_ECOLI
37	44	44.9	419	1	GSC_DROME
38	43.5	44.4	59	1	HPN_HELPY
39	43.5	44.4	98	1	2C_DICDI
40	43.5	44.4	507	1	ZNT1_HUMAN
41	43.5	44.4	688	1	ARK2_BOVIN
42	43.5	44.4	688	1	ARK2_HUMAN
43	43	43.9	335	1	HYPB_RHOCA
44	43	43.9	424	1	S3B4_HUMAN
45	42.5	43.4	689	1	ARK1_BOVIN

RESULT 1

KNG_HUMAN	STANDARD;	PRT;	644	AA.
ID	KNG_HUMAN			
AC	P01042; P01043;			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	01-FEB-1996 (Rel. 33, Last sequence update)			
DT	10-OCT-2003 (Rel. 42, Last annotation update)			
DE	Kinogen precursor (Alpha-2-thiol proteinase inhibitor) [Contains: Bradykinin].			
DE	Bradykinin].			
GN	KNG.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A. (ISOFORMS HMW AND LMW).			
RC	TISSUE=Liver;			
RX	MEDLINE=85234582; PubMed=2989293;			
RA	Takagaki Y., Kitamura N., Nakanishi S.;			
RT	"Cloning and sequence analysis of cDNAs for human high molecular weight and low molecular weight prekininogens. Primary structures of two human prekininogens.";			
RL	J. Biol. Chem. 260:8601-8609(1985).			
RN	[2]			
RP	GENE STRUCTURE.			
RX	MEDLINE=85234583; PubMed=2989294;			
RA	Kitamura N., Kitagawa H., Fukushima D., Takagaki Y., Miyata T.,			
RA	Nakanishi S.;			
RT	"Structural organization of the human kininogen gene and a model for its evolution.";			
RL	J. Biol. Chem. 260:8610-8617(1985).			
RN	[3]			
RP	SEQUENCE OF 1-401 FROM N.A.			
RX	MEDLINE=8512621; PubMed=6441591;			
RA	Okubo I., Kurachi K., Takasawa T., Shiokawa H., Sasaki M.;			
RT	"Isolation of a human cDNA for alpha 2-thiol proteinase inhibitor and its identity with low molecular weight kininogen.";			
RL	Biochemistry 23:5691-5697(1984).			
RN	[4]			
RP	SEQUENCE OF 379-644.			
RX	MEDLINE=86030270; PubMed=4054110;			
RA	Lottspeich F., Kellermann J., Henschen A., Foersts B.,			
RA	Mueller-Esterl W.;			
RT	"The amino acid sequence of the light chain of human high-molecular-mass kininogen.";			
RL	Eur. J. Biochem. 152:307-314(1985).			
RN	[5]			
RP	SEQUENCE OF 381-389.			
RX	MEDLINE=90255622; PubMed=4952632;			
RA	Pierce J.V.;			
RT	"Structural features of plasma kinins and kininogens.";			
RL	Fed. Proc. 27:52-57(1968).			
RN	[6]			
RP	DISULFIDE BONDS.			
RA	Sueyoshi T., Miyata T., Kato H., Iwanaga S.;			
RT	"Disulfide bonds in bovine HMW kininogens.";			

P17208 mus musculus
P10105 drosophila
P76425 escherichia
P54366 drosophila
Q48251 helicobacter
P15648 dictyosteli
Q9Y6M5 homo sapien
P36818 bos taurus
P35626 homo sapien
P26410 rhodobacter
Q15427 homo sapien
P21146 bos taurus

ALIGNMENTS

Seikagaku 56:808-808 (1984).

[7]

RE CARBOHYDRATE-LINKAGE SITE ASN-294.

RA MEDLINE=22660472; PubMed=12754519;

RT Zhang H., Li X.-J., Martin D.B., Abersold R.;

RT "Identification and quantification of N-linked glycoproteins using

RT hydrazide chemistry, stable isotope labeling and mass spectrometry."

Nat. Biotechnol. 21:660-666(2003).

CC -1- FUNCTION: (1) Kininogens are inhibitors of thiol proteases; (2)

CC HMW-kininogen plays an important role in blood coagulation by

CC helping to position optimally prekallikrein and factor XI next to

CC factor XII; (3) HMW-kininogen inhibits the thrombin- and plasmin-

CC induced aggregation of thrombocytes; (4) the active peptide

CC bradykinin that is released from HMW-kininogen shows a variety of

CC physiological effects: (4A) influence in smooth muscle

CC contraction, (4B) induction of hypotension, (4C) natriuresis and

CC diuresis, (4D) decrease in blood glucose level, (4E) it is a

CC mediator of inflammation and causes (4E1) increase in vascular

CC permeability, (4E2) stimulation of nociceptors (4E3) release of

CC other mediators of inflammation (e.g. prostaglandins), (4F) it has

CC a cardioprotective effect (directly via bradykinin action), (5)

CC indirectly via endothelium-derived relaxing factor action); (6)

CC LMW-kininogen inhibits the aggregation of thrombocytes; (6) LMW-

CC kininogen is in contrast to HMW-kininogen not involved in blood

CC clotting.

CC -1- SUBCELLULAR LOCATION: Secreted.

CC -1- ALTERNATIVE PRODUCTS:

CC Event=Alternative splicing; Named isoforms=2;

CC Name=HMW;

CC IsoId=P01042-1; Sequence=Displayed;

CC Name=LMW;

CC IsoId=P01042-2; Sequence=VSP_001261, VSP_001262;

CC -1- TISSUE SPECIFICITY: Plasma.

CC -1- PTM: Bradykinin is released from kininogen by plasma kallikrein.

CC -1- SIMILARITY: Contains 3 cystatin-like domains.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL; K02566; AAA35497.1; -.

DR EMBL; M11437; AAB59550.1; -.

DR EMBL; M11438; AAB59550.1; JOINED.

DR EMBL; M11521; AAB59550.1; JOINED.

DR EMBL; M11522; AAB59550.1; JOINED.

DR EMBL; M11523; AAB59550.1; JOINED.

DR EMBL; M11524; AAB59550.1; JOINED.

DR EMBL; M11525; AAB59550.1; JOINED.

DR EMBL; M11526; AAB59550.1; JOINED.

DR EMBL; M11527; AAB59550.1; JOINED.

DR EMBL; M11528; AAB59550.1; JOINED.

DR EMBL; M11437; AAB59551.1; -.

DR EMBL; M11438; AAB59551.1; JOINED.

DR EMBL; M11521; AAB59551.1; JOINED.

DR EMBL; M11522; AAB59551.1; JOINED.

DR EMBL; M11523; AAB59551.1; JOINED.

DR EMBL; M11524; AAB59551.1; JOINED.

DR EMBL; M11525; AAB59551.1; JOINED.

DR EMBL; M11526; AAB59551.1; JOINED.

DR EMBL; M11527; AAB59551.1; JOINED.

DR EMBL; M11528; AAB59551.1; JOINED.

DR PIR; A01279; KGHU1.

DR PIR; A01280; KGHU1.

DR SWISS-2DPAGE; P01042; HUMAN.

DR Genew; HGNC:6383; KNG.

DR MIM; 228960; -.

DR GO; GO:0007596; P: blood coagulation; NAS.

DR GO; GO:0030145; P: diuresis; NAS.

DR GO; GO:0005954; P: inflammatory response; NAS.

DR GO; GO:0030147; P: natriuresis; NAS.

DR GO; GO:006939; P: smooth muscle contraction; NAS.

DR InterPro; IPR00010; Cystatin.

DR InterPro; IPR002395; Kininogen.

DR Pfam; PF00031; cystatin; 3.

DR PRINTS; PR00334; Kininogen.

DR SMART; SM00043; CY; 3.

DR PROSITE; PS00287; CYSTATIN; 2.

CC Glycoprotein; Plasma; Repeat; Thiol protease inhibitor; Vasodilator;

CC Bradykinin; Blood coagulation; Inflammatory response; Signal;

CC Alternative splicing; Pyrrolidone carboxylic acid.

FT SIGNAL 1 18

FT CHAIN 19 644 KININOGEN.

FT CHAIN 19 380 KININOGEN HEAVY CHAIN.

FT PEPTIDE 381 389 BRADYKININ.

FT CHAIN 390 644 KININOGEN LIGHT CHAIN.

FT CHAIN 390 644 KININOGEN-LIKE 1.

FT DOMAIN 19 136 CYSTATIN-LIKE 2.

FT DOMAIN 137 258 CYSTATIN-LIKE 3.

FT DOMAIN 259 380 HIS-RICH

FT DOMAIN 420 510 (ASSOCIATED WITH CLOTTING ACTIVITY).

FT REPEAT 420 449

FT REPEAT 450 479

FT REPEAT 480 510

FT MOD RES 19 19

FT DISULFID 28 614

FT DISULFID 83 94

FT DISULFID 107 126

FT DISULFID 142 145

FT DISULFID 206 218

FT DISULFID 229 248

FT DISULFID 264 267

FT DISULFID 328 340

FT DISULFID 351 370

FT CARBOHYD 48 48

FT CARBOHYD 169 169

FT CARBOHYD 205 205

FT CARBOHYD 294 294

FT CARBOHYD 401 401

FT CARBOHYD 533 533

FT CARBOHYD 542 542

FT CARBOHYD 546 546

FT CARBOHYD 557 557

FT CARBOHYD 571 571

FT CARBOHYD 577 577

FT CARBOHYD 593 593

FT CARBOHYD 628 628

FT VARSPIC 402 427

FT VARSPIC 428 644

FT CONFLICT 593 593

FT SEQUENCE 644 AA; 71945 MW; 3132B4CBAP8FB7E CRC64;

Query Match 98.0%; Score 96; DB 1; Length 644;

Best Local Similarity 94.1%; Pred. No. 2.8e-07;

Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 HGHEQOHCGLGHGKXKL 17

DB 463 HGHEQOHCGLGHGKXKL 479

RESULT 2

KNH2_BOVIN

ID KNH2_BOVIN STANDARD; PRT; 619 AA.

AC P01045;

DT 21-JUL-1986 (Rel. 01, Created)

DT 21-JUL-1986 (Rel. 01, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Kininogen, HMW II precursor (Thiol proteinase inhibitor) [Contains:

DE Bradykinin].

Fri Mar 5 08:53:29 2004

"A single gene for bovine high molecular weight and low molecular weight kininogens.";
 Nature 305:545-549(1983).
 [2]
 RP MEDLINE=97137530; PubMed=3546295;
 RA Sueyoshi T., Miyata T., Hashimoto N., Kato H., Hayashida H.,
 Miyata T., Iwanaga S.;
 RA "Bovine high molecular weight kininogen. The amino acid sequence, positions of carbohydrate chains and disulfide bridges in the heavy chain portion.";
 J. Biol. Chem. 262:2768-2779(1987).
 [3]
 RP MEDLINE=70180420; PubMed=4986212;
 RA Kato H., Nagasawa S., Suzuki T.;
 RA "Studies on the structure of bovine kininogen: cleavages of disulfide bonds and of methionyl bonds in kininogen-II.";
 J. Biochem. 67:313-323(1970).
 [4]
 RP MEDLINE=75170265; PubMed=1169237;
 RA Han Y.N., Komiya M., Iwanaga S., Suzuki T.;
 RA "Studies on the primary structure of bovine high-molecular-weight kininogen. Amino acid sequence of a fragment ('histidine-rich peptide') released by plasma kallikrein.";
 J. Biochem. 77:55-68(1975).
 CC -!- FUNCTION: (1) Kininogens are inhibitors of thiol proteases; (2) HMW-kininogen plays an important role in blood coagulation by helping to position optically prekallikrein and factor XI next to factor XII; (3) HMW-kininogen inhibits the thrombin- and plasmin-induced aggregation of thrombocytes; (4) the active peptide bradykinin that is released from HMW-kininogen shows a variety of physiological effects: (4A) influence in smooth muscle contraction, (4B) induction of hypotension, (4C) natriuresis and diuresis, (4D) decrease in blood glucose level, (4E) it is a mediator of inflammation and causes (4E1) increase in vascular permeability, (4E2) stimulation of nociceptors (4E3) release of other mediators of inflammation (e.g. prostaglandins), (4F) it has a cardioprotective effect (directly via bradykinin action, indirectly via endothelium-derived relaxing factor action).
 CC -!- SUBCELLULAR LOCATION: Extracellular.
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event-Alternative splicing; Named isoforms=2;
 CC Name=HMW I;
 CC IsoId=P01044-1; Sequence=Displayed;
 CC Name=LMW I;
 CC IsoId=P01046-1; Sequence=External;
 CC -!- TISSUE SPECIFICITY: Plasma.
 CC -!- PTM: Bradykinin is released from kininogen by plasma kallikrein.
 CC -!- SIMILARITY: Contains 3 cystatin-like domains.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

 DR EMBL: V01491; CAA24735.1; -;
 DR PIR: A01281; KGB0H1.
 DR InterPro: IPR000010; Cystatin.
 DR InterPro: IPR002395; Kininogen.
 DR Pfam: PF00031; cystatin; 3.
 DR PRINTS: PR00334; KININOGEN.
 DR SMART: SM00043; CY; 3.
 DR PROSITE: PS00287; CYSSTATIN; 2.
 KW Glycoprotein; Plasma; Repeat; Vasodilator; Alternative splicing;
 KW Thiol protease inhibitor; Bradykinin; Blood coagulation;
 KW Inflammatory response; Signal; Pyrrolidone carboxylic acid.
 FT SIGNAL 1 18 PROBABLE.
 FT CHAIN 19 621 KININOGEN, HMW I.

FT CHAIN 19 378 HEAVY CHAIN.
 FT PEPTIDE 380 388 BRADYKININ.
 FT CHAIN 389 621 LIGHT CHAIN.
 FT DOMAIN 19 135 CYSTATIN-LIKE 1.
 FT DOMAIN 136 257 CYSTATIN-LIKE 2.
 FT DOMAIN 258 378 CYSTATIN-LIKE 3.
 FT MOD RES 19 19 PYRROLIDONE CARBOXYLIC ACID.
 FT CARBOHYD 87 87 N-LINKED (GLCNAC. . .).
 FT CARBOHYD 136 136 O-LINKED (PARTIAL).
 FT CARBOHYD 168 168 N-LINKED (GLCNAC. . .) (OR 169).
 FT CARBOHYD 197 197 N-LINKED (GLCNAC. . .) (PARTIAL).
 FT CARBOHYD 204 204 N-LINKED (GLCNAC. . .).
 FT DISULFID 27 591 INTERCHAIN.
 FT DISULFID 82 93
 FT DISULFID 106 125
 FT DISULFID 141 144
 FT DISULFID 205 217
 FT DISULFID 228 247
 FT DISULFID 263 266
 FT DISULFID 327 339
 FT DISULFID 350 369
 SQ SEQUENCE 621 AA; 68890 MW; D16850BEFE3C55CD CRC64;
 Query Match 82.7%; Score 81; DB 1; Length 621;
 Best Local Similarity 85.7%; Pred. No. 5,1e-05;
 Matches 12; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 1 HGHEQHQHGLGHGHEK 14
 DB 463 HGHEQHQHGLGHGHEK 476
 RESULT 4
 KNG RAT STANDARD; PRT; 639 AA.
 ID_KNG RAT
 AC P08934; P08933;
 DT 01-NOV-1988 (Rel. 09, Created)
 DT 01-NOV-1988 (Rel. 09, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Kininogen precursor [Contains: Bradykinin].
 GN KNG.
 OS Rattus norvegicus (Rat).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 CC NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORMS HMW AND LMW).
 RX MEDLINE=87137443; PubMed=3029068;
 RA Kitagawa H., Kitamura N., Hayashida H., Miyata T., Nakanishi S.;
 RT "Differing expression patterns and evolution of the rat kininogen gene family.";
 J. Biol. Chem. 262:2190-2198(1987).
 [2]
 RP SEQUENCE FROM N.A. (ISOFORM LMW).
 RX MEDLINE=86008264; PubMed=2413018;
 RA Furuto-Kato S., Matsumoto A., Kitamura N., Nakanishi S.;
 RT "Primary structures of the mRNAs encoding the rat precursors for bradykinin and T-kinin. Structural relationship of kininogens with major acute phase protein and alpha 1-cysteine proteinase inhibitor.";
 J. Biol. Chem. 260:12054-12059(1985).
 [3]
 RP SEQUENCE OF 1-65 FROM N.A.
 RC STRAIN=Buffalo;
 RX MEDLINE=87250580; PubMed=2439509;
 RA Fung W.-P., Schreiber G.;
 RT "Structure and expression of the genes for major acute phase alpha 1-protein (thiostatin) and kininogen in the rat.";
 J. Biol. Chem. 262:9298-9308(1987).
 [4]
 RP SEQUENCE OF 1-41 FROM N.A.
 RC STRAIN=Wistar; TISSUE=Liver;
 RX MEDLINE=87137465; PubMed=3818598;

RA Kageyama R., Kitamura N., Okubo H., Nakanishi S.;
 RT "Differing utilization of homologous transcription initiation sites
 RL J. Biol. Chem. 262:2345-2351(1987).
 CC -!- FUNCTION: (1) kininogens are inhibitors of thiol proteases; (2)
 CC HMW-kininogen plays an important role in blood coagulation by
 CC helping to position optimally prekallikrein and factor XI next to
 CC factor XII; (3) HMW-kininogen inhibits the thrombin- and plasmin-
 CC induced aggregation of thrombocytes; (4) the active peptide
 CC bradykinin that is released from HMW-kininogen shows a variety of
 CC physiological effects: (4A) influence in smooth muscle
 CC contraction, (4B) induction of hypotension, (4C) natriuresis and
 CC diuresis, (4D) decrease in blood glucose level, (4E) it is a
 CC mediator of inflammation and causes (4E1) increase in vascular
 CC permeability, (4E2) stimulation of nociceptors (4E3) release of
 CC other mediators of inflammation (e.g. prostaglandins), (4F) it has
 CC a cardioprotective effect (directly via bradykinin action,
 CC indirectly via endothelium-derived relaxing factor action); (5)
 CC LMW-kininogen inhibits the aggregation of thrombocytes; (6) LMW-
 CC kininogen is in contrast to HMW-kininogen not involved in blood
 CC clotting.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=HMW;
 CC IsoId=P08934-1; Sequences=Displayed;
 CC Name=LMW;
 CC IsoId=P08934-2; Sequences=VSP_001265, VSP_001266;
 CC -!- TISSUE SPECIFICITY: Plasma.
 CC -!- PTM: Bradykinin is released from kininogen by plasma kallikrein.
 CC -!- MISCELLANEOUS: Rats express four types of kininogens: the
 CC classical HMW/LMW kininogens and two additional LMW-like
 CC kininogens: T-I and T-II.
 CC -!- SIMILARITY: Contains 3 cystatin-like domains.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; L29428; AAA41486.1; -.
 DR EMBL; M1684; AAA41487.1; -.
 DR EMBL; M14369; AAA41484.1; -.
 DR EMBL; M14369; AAA41485.1; ALT_SEQ.
 DR EMBL; M16455; AAA41482.1; -.
 DR PIR; A25486; A25486.
 DR PIR; A28055; A28055.
 DR InterPro; IPR000010; Cystatin.
 DR InterPro; IPR002395; Kininogen.
 DR Pfam; PF00031; cystatin; 3.
 DR PRINTS; PR00334; KININOGEN.
 DR SMART; SM00043; Cy; 3.
 DR PROSITE; PS00287; CYSTATIN; 2.
 KW Glycoprotein; Plasma; Repeat; Thiol protease inhibitor; Vasodilator;
 KW Bradykinin; Blood coagulation; Inflammatory response; Signal;
 KW Alternative splicing; Multigene family.
 FT SIGNAL 1 18
 FT CHAIN 19 639 KININOGEN.
 FT CHAIN 19 380 KININOGEN HEAVY CHAIN.
 FT PEPTIDE 381 389 BRADYKININ.
 FT CHAIN 390 639 KININOGEN LIGHT CHAIN.
 FT DOMAIN 19 136 CYSTATIN-LIKE 1.
 FT DOMAIN 137 258 CYSTATIN-LIKE 2.
 FT DOMAIN 259 380 CYSTATIN-LIKE 3.
 FT DOMAIN 439 514 HIS-RICH.
 FT DISULFID 28 609 INTERCHAIN (BY SIMILARITY).
 FT DISULFID 83 94 BY SIMILARITY.
 FT DISULFID 107 126 BY SIMILARITY.
 FT DISULFID 142 145 BY SIMILARITY.
 FT DISULFID 206 218 BY SIMILARITY.

FT DISULFID 229 248 BY SIMILARITY.
 FT DISULFID 264 267 BY SIMILARITY.
 FT DISULFID 328 340 BY SIMILARITY.
 FT DISULFID 351 370 BY SIMILARITY.
 FT CARBOHYD 82 82 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 127 127 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 169 169 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 205 205 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 294 294 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 529 529 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARSPLIC 402 433 VSPSTVRIQVERDPNGQPIHGHWLHAQ -> RLNS
 FT CRYXRLKAGAPAPQAEASTVTP (in isoform
 FT LMW).
 FT /FTID=VSP_001265.
 FT VARSPLIC 434 639 Missing (in isoform LMW).
 FT /FTID=VSP_001266.
 FT CONFLICT 61 61 E > K (IN REF. 2).
 SQ SEQUENCE 639 AA; 70933 MW; D317DF94FF56AF5 CRC64;
 Query Match 80.6%; Score 79; DB 1; Length 639;
 Best Local Similarity 70.6%; Pred. No. 0.00011;
 Matches 12; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
 QY 1 HGHEQOHGLGHGHXXKL 17
 DB 450 HGHEQOHGLGHGHXXKL 466
 RESULT 5
 KMG MOUSE
 ID KMG MOUSE STANDARD; PRT; 661 AA.
 AC 008677; 008676; Q91XK5;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Kininogen precursor [Contains: Bradykinin].
 GN KMG.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORMS HMW AND LMW).
 RC STRAIN=C57BL/6 X CBA; TISSUE=Liver;
 RX MEDLINE=97342556; PubMed=9199253;
 RA Takano M., Kondo J., Yavama K., Okani M., Sano K., Okamoto H.;
 RA Baldrilli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
 RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
 RA Blake J.A., Bradt D., Brueic V., Chothia C., Corbani L.E., Cousins S.,
 RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
 RA Gasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
 RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
 RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
 RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
 RA Maglott D.R., Maltails L., Marchionni L., McKenzie L., Miki H.,
 RA Nagashima T., Nunata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
 RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
 RA Ravasi T., Reed J.C., Reed D.J., Reid G., Ring B.Z., Ringwald M.,
 RA Saitanin A., Schneider C., Semple C.A., Setou M., Shimada K.,
 RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
 RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
 RA Wilmink L.G., Wynshaw-Boris A., Yangisawa M., Yang I., Yang L.,
 RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,

RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume M., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sakai D., Shibata K., Shinagawa A.,
RA Yasunishi A., Yoshino K., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.,
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs";
RL Nature 420:563-573 (2002).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM LMW).
RC TISSUE=Liver;
RX MEDLINE=22398257; PubMed=12477932;
RA Strausberg R.L., Reingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altshul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Lequellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Guntaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Pahey J., Helton E., Kettner M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.N.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16999-16903 (2002).
CC -!- FUNCTION: (1) Kininogens are inhibitors of thiol proteases; (2)
CC HMW-kininogen plays an important role in blood coagulation by
CC helping to position optimally prekallikrein and factor XI next to
CC factor XII; (3) HMW-kininogen inhibits the thrombin- and plasmin-
CC induced aggregation of thrombocytes; (4) the active peptide
CC bradykinin that is released from HMW-kininogen shows a variety of
CC physiological effects: (4A) influence in smooth muscle
CC contraction, (4B) induction of hypertension, (4C) natriuresis and
CC diuresis, (4D) decrease in blood glucose level, (4E) it is a
CC mediator of inflammation and causes (4E1) increase in vascular
CC permeability, (4E2) stimulation of nociceptors (4E3) release of
CC other mediators of inflammation (e.g. prostaglandins), (4F) it has
CC a cardioprotective effect (directly via bradykinin action,
CC indirectly via endothelium-derived relaxing factor action); (5)
CC LMW-kininogen inhibits the aggregation of thrombocytes; (6) LMW-
CC kininogen is in contrast to HMW-kininogen not involved in blood
CC clotting (by similarity).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- ALTERNATIVE PRODUCTS:
CC Name=HMW;
CC IsoId=O08677-1; Sequence=Displayed;
CC Name=LMW;
CC IsoId=O08677-2; Sequence=VSP_001263; VSP_001264;
CC -!- TISSUE SPECIFICITY: Plasma.
CC -!- PTM: Bradykinin is released from kininogen by plasma kallikrein.
CC -!- SIMILARITY: Contains 3 cystatin-like domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.ebi.ac.uk/announcements>
CC or send an email to license@ebi.ac.uk).
CC -----
CC EMBL; D84435; BAAL19743.1; -;
CC EMBL; D84415; BAAL19742.1; -;
CC EMBL; AK005547; BAB24115.1; -;
CC EMBL; BC018158; AAHL18158.1; -;

DR MGD; MGI:1097705; Kng.
DR InterPro; IPR000010; Cystatin.
DR InterPro; IPR002395; Kininogen.
DR Pfam; PF00031; cystatin; 3.
DR PRINTS; PR00334; KININOGEN.
DR SMART; SM00043; Cy; 3.
DR PROSITE; PS00287; CYSTATIN; 1.
DR GlycoProtein; Plasma; Repeat; Thiol protease inhibitor; Vasodilator;
DR Bradykinin; Blood coagulation; Inflammatory response; Signal;
KW Alternative splicing.
FT SIGNAL 1 18 POTENTIAL.
FT CHAIN 19 661 KININOGEN.
FT CHAIN 19 379 KININOGEN HEAVY CHAIN.
FT CHAIN 380 388 BRADYKININ.
FT CHAIN 389 661 KININOGEN LIGHT CHAIN.
FT DOMAIN 19 135 CYSTATIN-LIKE 1.
FT DOMAIN 136 257 CYSTATIN-LIKE 2.
FT DOMAIN 258 379 CYSTATIN-LIKE 3.
FT DOMAIN 439 524 HIS-RICH.
FT DISULFID 28 631 INTERCHAIN (BY SIMILARITY).
FT DISULFID 83 94 BY SIMILARITY.
FT DISULFID 107 125 BY SIMILARITY.
FT DISULFID 141 144 BY SIMILARITY.
FT DISULFID 205 217 BY SIMILARITY.
FT DISULFID 228 247 BY SIMILARITY.
FT DISULFID 263 266 BY SIMILARITY.
FT DISULFID 327 339 BY SIMILARITY.
FT DISULFID 350 369 BY SIMILARITY.
FT CARBOHYD 82 82 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 168 168 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 204 204 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 242 242 N-LINKED (GLCNAC. .) (POTENTIAL).
FT VARSPLIC 401 432 VSPPIYARLQEEEDAEQCGTHGHWLHEKQ -> RLIRA
CEYGRSLKAGAPAPAPQAESSQVKQ (in isoform
LMW).
FT FT /FTid=VSP_001263.
FT VARSPLIC 433 661 Missing (in isoform LMW).
FT /FTid=VSP_001264.
SQ SEQUENCE 661 AA; 73102 MW; 774460258D58796E CRC64;
Query Match 80.6%; Score 79; DB 1; Length 661;
Best Local Similarity 70.6%; Pred. No. 0.00011;
Matches 12; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
QY 1 HGHEQHQHGLGHGXKL 17
Db 460 HGHEQHQHGLGHGXKL 476
|||::|||::|||::|||
|||::|||::|||::|||
RESULT 6
CAUP_DROME
ID CAUP_DROME STANDARD; PRT; 693 AA.
AC P54269; Q9VU00;
DT 01-OCT-1996 (Rel. 34, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Homeobox protein caupolican.
GN CAUP OR CG10605.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE=96180722; PubMed=8620542;
RA Gomez-Skarmeta J.-L., del Corral R.D., de la Calle-Mustienes E.,
RA Ferrer-Marco D., Modolell J.;
RT "Araucan and caupolican, two members of the novel ironouls complex,
RL encode homeoproteins that control proneural and vein-forming genes.";
RN Cell 85:95-110 (1996).
RP SEQUENCE FROM N.A.

RC STRAIN=Berkeley;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazef R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abriel J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.T., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Borkan M.R., Bouck J., Brokstein P., Brotter P.,
 RA Burks K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabrielle A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Talali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shier B.C., Siden-Kiamos I., Stamps M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirska R., Tector R., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RA "The genome sequence of *Drosophila melanogaster*.";
 BL Science 287:2185-2195 (2000).
 CC -!- FUNCTION: Controls proneural and vein forming genes. Positive
 CC transcriptional control of AC-SC (achaete-scute). May act as an
 CC activator that interacts with the transcriptional complex
 CC assembled on the AC and SC promoters and participates in
 CC transcription initiation.
 CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -!- MISCELLANEOUS: 'Cupolican' is named after the Araucanian
 CC medial stripe of hairs on the head.
 CC -!- SIMILARITY: Belongs to the TALE/IRO homeobox family.
 CC -!- SIMILARITY: Contains 1 homeobox domain.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; X95178; CAA64485.1; -;
 CC EMBL; AB003540; AAF49895.1; -;
 CC HSP; P41778; 1D06.
 CC FlyBase; FBgn015919; caup.
 CC InterPro; IPR001356; Homeobox.
 CC InterPro; IPR003893; Iroquois/homeo.
 CC Pfam; PF00046; homeobox; 1.
 CC ProDom; PD000010; Homeobox; 1.
 CC SMART; SM00389; Hox; 1.
 CC SMART; SM00548; IRO; 1.
 CC PROSITE; PS00027; HOMEBOX_1; 1.
 CC PROSITE; PS00071; HOMEBOX_2; 1.
 CC Transcription regulation; Activator; DNA-binding; Homeobox;
 KW

KW Nuclear protein; Developmental protein.
 FT DNA_BIND 226 288 HOMEBOX (TALE-TYPE).
 FT DOMAIN 300 303 POLY-ASP.
 FT DOMAIN 405 418 POLY-GLN.
 FT DOMAIN 501 516 POLY-GLN.
 FT DOMAIN 517 528 POLY-HIS.
 FT DOMAIN 565 572 POLY-SER.
 FT DOMAIN 613 624 POLY-SER.
 FT CONFLICT 106 106 C -> R (IN REF. 1).
 FT CONFLICT 316 316 G -> A (IN REF. 1).
 FT CONFLICT 678 678 G -> A (IN REF. 1).
 SQ SEQUENCE 693 AA; 73667 MW; FEEB1616493F7EC9 CFC64;
 Query Match 65.3%; Score 64; DB 1; Length 693;
 Best Local Similarity 76.9%; Pred. No. 0.022;
 Matches 10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 HGHFQHQHGLGHGH 13
 DB 656 HGHGHGHGLGHGH 668
 RESULT 7
 S109_RABIT STANDARD; PRT; 118 AA.
 AC P50117;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Calgranulin B (Migration inhibitory factor-related protein 14)
 DE (MRP-14) (Fragment).
 GN S100A9 OR MRP-14.
 OS Eucaryotagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 OX NCBI_TaxID=9986;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=New Zealand white; TISSUE=Neutrophils;
 RX MEDLINE=96355278; PubMed=8702688;
 RA Yang Z., de Veer M.J., Gardiner E.E., Devenish R.J., Handley C.J.,
 RA Underwood J.R., Robinson H.C.;
 RT "Rabbit polymorphonuclear neutrophils form 35S-labeled S-sulfo-
 RT calgranulin C when incubated with inorganic [35S] sulfate.";
 RL J. Biol. Chem. 271:19802-19809 (1996).
 RN [2]
 RP SEQUENCE OF 45-82 FROM N.A.
 RC STRAIN=New Zealand white;
 RX MEDLINE=94198229; PubMed=8148323;
 RA Mori S., Goto K., Goto F., Matakami K., Ohkawara S., Yoshinaga M.;
 RT "Dynamic changes in mRNA expression of neutrophils during the course
 RT of acute inflammation in rabbits.";
 RL Int. Immunol. 6:149-156 (1994).
 CC -!- SIMILARITY: Belongs to the S-100 family.
 CC -!- SIMILARITY: Contains 2 EF-hand calcium-binding domains.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AF091849; AAC61771.1; -;
 CC EMBL; D17404; BAA04227.1; -;
 CC PIR; I46861; I46861.
 CC HSP; P80511; IESB.
 CC InterPro; IPR001751; CaBP_S100.
 CC InterPro; IPR002048; EF-hand.
 CC Pfam; PF00036; efhand; 1.
 CC Pfam; PF01023; S100; 1.
 CC ProDom; PD003407; CaBP_S100; 1.


```

DR PRODOM; PD000012; EF-hand; 1.
DR PROSITE; PS00018; EF HAND; 1.
DR PROSITE; PS00303; S100 CABP; 1.
KW Calcium-binding; Repeat.
FT NON TER 1
FT CA_BIND 9 22 EF-HAND 1 (LOW AFFINITY) (POTENTIAL).
FT CA_BIND 53 64 EF-HAND 2 (HIGH AFFINITY) (POTENTIAL).
FT DOMAIN 103 118 2 X 8 AA TANDEM REPEATS OF G-H-G-H-
G-H-S-H.
FT REPEAT 103 110 1.
FT REPEAT 111 118 2.
SQ SEQUENCE 118 AA; 13292 MW; 7496118E21AD5041 CRC64;

Query Match 60.2%; Score 59; DB 1; Length 118;
Best Local Similarity 69.2%; Pred. No. 0.02; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 4;

QY 1 HGHEQHQHGLGHGH 13
DB 104 HGHGSHGHGHGH 116

RESULT B
VG50 HSV11
ID VG50 HSV11 STANDARD; PRT; 670 AA.
AC Q00130;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Hypothetical gene 50 protein.
GN 50.
OS Ictalurid herpesvirus 1 (Channel catfish virus) (CCV).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Ictalurid Herpes-like viruses.
OX NCBI_TaxID=10401;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=Alburn 1.
RX MEDLINE=92087490; PubMed=1727613;
RA Davison A.J.;
RL "Channel catfish virus: a new type of herpesvirus.";
RT Virology 186:9-14(1992).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M75136; AAA88153.1; -.
CC PIR; F36791; F36791.
CC KW Hypothetical protein; Repeat.
FT REPEAT 143 158 1.
FT REPEAT 171 186 2.
FT REPEAT 200 214 3.
FT REPEAT 215 233 4.
FT REPEAT 234 252 5.
FT REPEAT 253 268 6.
FT REPEAT 279 293 7.
FT REPEAT 294 309 8.
FT REPEAT 320 334 9.
FT REPEAT 335 349 10.
FT REPEAT 362 376 11.
FT REPEAT 377 391 12.
FT REPEAT 392 406 13.
FT REPEAT 407 421 14.
FT REPEAT 422 436 15.
FT REPEAT 437 452 16.
FT REPEAT 464 477 17.
FT REPEAT 478 493 18.
FT REPEAT 504 517 19.

```

```

FT REPEAT 518 531 20.
FT REPEAT 532 545 21.
FT REPEAT 546 559 22.
FT REPEAT 560 573 23.
FT REPEAT 574 587 24.
FT REPEAT 588 601 25.
FT REPEAT 602 615 26.
FT REPEAT 616 629 27.
SQ SEQUENCE 670 AA; 64174 MW; 2B64A781C519E8B4 CRC64;

Query Match 58.2%; Score 57; DB 1; Length 670;
Best Local Similarity 69.2%; Pred. No. 0.25; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 4;

QY 1 HGHEQHQHGLGHGH 13
DB 637 HGHGSHGHGHGH 649

RESULT 9
FSH DROME
ID FSH DROME STANDARD; PRT; 2038 AA.
AC P13709; P13710;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Female sterile homeotic protein (Fragile-chorion membrane protein).
GN FS(1)H OR FSH.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Phyllophaga; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89276730; PubMed=2567251;
RA Haynes S.R., Moser B.A., Bhatia-Dey N., Dawid I.B.;
RT "The Drosophila fsh locus, a maternal effect homeotic gene, encodes
RT apparent membrane proteins.";
RL Dev. Biol. 134:246-257(1989).
CC -I- FUNCTION: Required maternally for proper expression of other
CC homeotic genes involved in pattern formation, such as UBX.
CC -I- SIMILARITY: HIGH, TO HUMAN RING3 PROTEIN.
CC -I- SIMILARITY: Contains 2 bromodomains.
CC -I- SIMILARITY: Contains 1 ET domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M23221; AAA28540.1; -.
CC EMBL; M23222; AAA28541.1; ALT_TERM.
CC EMBL; M15762; AAA70424.1; -.
CC EMBL; M15763; AAA70423.1; -.
CC EMBL; M15764; AAA70422.1; -.
CC PIR; A43742; A43742.
CC HSSP; Q92831; 1B91.
CC Flybase; FBgn0004656; fs(1)h.
CC InterPro; IPR001487; Bromodomain.
CC Pfam; PF00439; bromodomain; 2.
CC PRINTS; PR00503; BROMODOMAIN.
CC SMART; SM00297; BRNMC; 2.
CC PROSITE; PS00633; BROMODOMAIN_1; 2.
CC PROSITE; PS0014; BROMODOMAIN_2; 2.
CC Developmental protein; Bromodomain; Transmembrane; Repeat.
FT DOMAIN 51 123 BROMODOMAIN 1.
FT DOMAIN 495 567 BROMODOMAIN 2.
FT DOMAIN 945 1106 ET DOMAIN.
FT TRANSMEM 330 350 POTENTIAL.

```

```

FT TRANSMEM 451 471 POTENTIAL.
FT TRANSMEM 750 770 POTENTIAL.
FT TRANSMEM 790 810 POTENTIAL.
FT TRANSMEM 816 830 POTENTIAL.
FT TRANSMEM 874 894 POTENTIAL.
FT TRANSMEM 1731 1751 POTENTIAL.
FT TRANSMEM 1939 1959 POTENTIAL.
FT VARIANT 909 909 G -> A.
FT VARIANT 1022 1022 H -> RKPYY.
FT VARIANT 2038 AA; 205332 MW; 849E0706D50A0098 CRC64;
SQ SEQUENCE 2038 AA; 205332 MW; 849E0706D50A0098 CRC64;

Query Match 58.2%; Score 57; DB 1; Length 2038;
Best Local Similarity 69.2%; Pred. No. 0.78;
Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 HGHEQOQHGLGHGH 13
DB 596 HGHEQOQHGLGHGH 608

RESULT 10
COTL_YEAST
ID COTL_YEAST STANDARD; PRT; 439 AA.
AC P32798;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Cobalt uptake protein COTL.
GN COTL OR YOR316C OR O6131.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c / FY1679;
RX MEDLINE=92375034; PubMed=1508175;
RA Conklin D.S., McMaster J.A., Culbertson M.R., Kung C.;
RT "COTL, a gene involved in cobalt accumulation in Saccharomycetes cerevisiae."
RL Mol. Cell. Biol. 12:3678-3688 (1992).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c / FY1679;
RX MEDLINE=97051589; PubMed=896266;
RA Pearson B.W., Hernandez Y., Payne J., Wolf S.S., Kalogeropoulos A.,
RT Schweizer M.;
RT "Sequencing of a 35.71 kb DNA segment on the right arm of yeast chromosome XV reveals regions of similarity to chromosomes I and XIII."
RL Yeast 12:1021-1031 (1996).
CC -1- FUNCTION: PROBABLY RESPONSIBLE FOR THE UPTAKE OF COBALT IONS. IT APPEARS TO ACT IN A DOSAGE-DEPENDENT MANNER TO COUNTERACT THE ADVERSE EFFECTS OF COBALT IONS ON CELLS. IT MAY PARTICIPATE IN THE REGULATION OF COBALT LEVELS UNDER NORMAL PHYSIOLOGICAL CONDITIONS AND MAY BE IMPORTANT IN THE SUPPLY OF METAL THAT IS REQUIRED FOR METALLOENZYME OR COFACTOR SYNTHESIS. IT REDUCES THE TOXICITY OF COBALT AND RHODIUM IONS. OTHER COMPONENTS RESPONSIBLE FOR COBALT TRANSPORT EXIST.
CC -1- SUBCELLULAR LOCATION: MITOCHONDRIAL MEMBRANE. ANOTHER POSSIBILITY EXISTS THAT IT IS ASSOCIATED WITH ANOTHER UNIDENTIFIED MEMBRANE THAT HAS BEEN ENRICHED IN THE MITOCHONDRIAL MEMBRANE FRACTIONS.
CC -1- SIMILARITY: Belongs to the cation diffusion facilitator family of transporters (GDF, TC 2.A.4). SLC30A subfamily.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
CC EMBL; N88252; AAA74884.1; -.

```

```

DR EMBL; X90565; CAA62171.1; -.
DR EMBL; Z75224; CAA9636.1; -.
DR PIR; S58327; S58327.
DR GenOnline; 143904; -.
DR SGD; S0005843; COTL.
DR GO; GO:0000324; C:cytosome (sensu Fungi); IDA.
DR GO; GO:0015087; F:cobalt ion transporter activity; IMP.
DR GO; GO:0005385; F:zinc ion transporter activity; IGI.
DR GO; GO:0006824; P:cobalt ion transport; IMP.
DR GO; GO:0006822; P:zinc ion homeostasis; IMP.
DR GO; GO:0006829; P:zinc ion transport; IGI.
DR InterPro; IPR002524; Cation efflux; 1.
DR TIGRFAMs; TIGR01297; CDF; 1.
KW TRANSPORT; Cobalt; Mitochondrion; Transmembrane.
FT TRANSMEM 10 27 POTENTIAL.
FT TRANSMEM 43 60 POTENTIAL.
FT TRANSMEM 78 100 POTENTIAL.
FT TRANSMEM 114 133 POTENTIAL.
FT TRANSMEM 244 265 POTENTIAL.
FT TRANSMEM 279 295 POTENTIAL.
FT DOMAIN 140 148 HIS-RICH; COULD BE INVOLVED IN COORDINATION OF COBALT IONS.
FT DOMAIN 163 169 HIS-RICH; COULD BE INVOLVED IN COORDINATION OF COBALT IONS.
FT CONFLICT 227 227 G -> E (IN REF. 1).
FT CONFLICT 333 334 HI -> RV (IN REF. 1).
FT CONFLICT 424 424 I -> V (IN REF. 1).
SQ SEQUENCE 439 AA; 48154 MW; AC88AAAF2EE4AED CRC64;

Query Match 57.1%; Score 56; DB 1; Length 439;
Best Local Similarity 61.5%; Pred. No. 0.23;
Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 HGHEQOQHGLGHGH 13
DB 134 HDNDQEHGHGHGH 146

RESULT 11
ZNTL_MOUSE
ID ZNTL_MOUSE STANDARD; PRT; 503 AA.
AC Q60738;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Zinc transporter 1 (Znt-1).
GN SLC30A1 OR ZNT1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=95188868; PubMed=7882967;
RA Palmiter R.D., Findley S.D.;
RT "Cloning and functional characterization of a mammalian zinc transporter that confers resistance to zinc."
RL EMBO J. 14:639-649 (1995).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=22388257; PubMed=12477932;
RA Strauberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shemmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buelow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heien F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Uesdin T.B., Toshiyuki S., Carninci P., Prange C., Rana S.S., Iqbalano N.A., Peters G.J., Abramson R.D., Mullany S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

```

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton A., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
CC -!- FUNCTION: May be involved in zinc transport out of the cell.
CC Lethality of knockout early in gestation suggests a role of the
CC protein in fetal zinc acquisition and retention.
CC -!- SUBUNIT: Multimer (Probable).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Localized on the
CC plasma membrane (Probable).
CC -!- TISSUE SPECIFICITY: Widely expressed.
CC -!- SIMILARITY: Belongs to the cation diffusion facilitator family of
CC transporters (CDF, TC 2.A.4). SLC30A subfamily.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; U17132; AA79233.1; -.
CC EMBL; BC052166; AAH52166.1; -.
CC PIR; S54302; S54302.
CC MGI; MGI:1345281; SLC30a1.
CC InterPro: IPR002524; Cation efflux.
CC Pfam; PF01545; Cation efflux; 1.
CC TIGRFAMs; TIGR01297; CDF; 1.
CC Zinc transport; Transmembrane; Multigene family; Repeat.
CC KW DOMAIN 1 10 CYTOPLASMIC (POTENTIAL).
CC FT TRANSMEM 11 31 POTENTIAL.
CC FT DOMAIN 22 35 EXTRACELLULAR (POTENTIAL).
CC FT TRANSMEM 36 56 POTENTIAL.
CC FT DOMAIN 57 78 CYTOPLASMIC (POTENTIAL).
CC FT TRANSMEM 79 99 POTENTIAL.
CC FT DOMAIN 100 113 EXTRACELLULAR (POTENTIAL).
CC FT TRANSMEM 114 134 POTENTIAL.
CC FT DOMAIN 135 243 CYTOPLASMIC (POTENTIAL).
CC FT TRANSMEM 244 264 POTENTIAL.
CC FT DOMAIN 265 303 EXTRACELLULAR (POTENTIAL).
CC FT TRANSMEM 304 324 POTENTIAL.
CC FT DOMAIN 325 503 CYTOPLASMIC (POTENTIAL).
CC FT DOMAIN 145 156 6 X 2 AA APPROXIMATE REPEATS OF H-G.
CC FT CARBOHYD 294 294 N-LINKED (GLCNAC...) (POTENTIAL).
CC SQ SEQUENCE 503 AA; 54716 MW; 7C4FF93FC13CDA22 CRC64;

Query Match 55.1%; Score 54; DB 1; Length 503;
Best Local Similarity 60.0%; Pred. No. 0.53;
Matches 9; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 GHEQHQHGLGHGKXK 16
Db 145 GHGSHGHGHGHLAK 159

RESULT 12
ID ZNT1 RAT STANDARD; PRT; 507 AA.
AC Q62720;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Zinc transporter 1 (Znt-1).
GN SLC30A1 OR ZNT1.
OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=95188868; PubMed=7882967;
RA Palmer R.D., Findley S.D.;
RT "Cloning and functional characterization of a mammalian zinc
transporter that confers resistance to zinc.";
RL EMBO J. 14:639-649 (1995).
RN [2]
RP INDUCTION BY ZINC.
RC TISSUE=Intestine;
RX MEDLINE=98226729; PubMed=9560190;
RA McMahon R.J., Cousins R.J.;
RT "Regulation of the zinc transporter Znt-1 by dietary zinc.";
RL Proc. Natl. Acad. Sci. U.S.A. 95:4841-4846 (1998).
CC -!- FUNCTION: May be involved in zinc transport out of the cell.
CC -!- SUBUNIT: Multimer (Probable).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC LOCALIZED ON THE PLASMA MEMBRANE (PROBABLE). LOCALIZED ON THE
CC BASOLATERAL SURFACE OF THE ENTEROCYTES.
CC -!- TISSUE SPECIFICITY: WIDELY EXPRESSED. THE PROTEIN IS DETECTED IN
CC DUODENUM AND JEJUNUM BUT NOT IN ILEUM AND COLON.
CC -!- INDUCTION: SLIGHTLY BY ZINC IN THE INTESTINE, BUT NOT THE LIVER.
CC -!- SIMILARITY: Belongs to the cation diffusion facilitator family of
CC transporters (CDF, TC 2.A.4). SLC30A subfamily.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; U17133; AA79234.1; -.
CC PIR; S54303; S54303.
CC InterPro: IPR002524; Cation efflux.
CC Pfam; PF01545; Cation efflux; 1.
CC TIGRFAMs; TIGR01297; CDF; 1.
CC Zinc transport; Transmembrane; Multigene family; Repeat.
CC KW DOMAIN 1 10 CYTOPLASMIC (POTENTIAL).
CC FT TRANSMEM 11 31 POTENTIAL.
CC FT DOMAIN 32 35 EXTRACELLULAR (POTENTIAL).
CC FT TRANSMEM 36 56 POTENTIAL.
CC FT DOMAIN 57 78 CYTOPLASMIC (POTENTIAL).
CC FT TRANSMEM 79 99 POTENTIAL.
CC FT DOMAIN 100 113 EXTRACELLULAR (POTENTIAL).
CC FT TRANSMEM 114 134 POTENTIAL.
CC FT DOMAIN 135 247 CYTOPLASMIC (POTENTIAL).
CC FT TRANSMEM 248 268 POTENTIAL.
CC FT DOMAIN 269 307 EXTRACELLULAR (POTENTIAL).
CC FT TRANSMEM 308 328 POTENTIAL.
CC FT DOMAIN 329 507 CYTOPLASMIC (POTENTIAL).
CC FT DOMAIN 145 156 6 X 2 AA APPROXIMATE REPEATS OF H-G.
CC FT CARBOHYD 298 298 N-LINKED (GLCNAC...) (POTENTIAL).
CC SQ SEQUENCE 507 AA; 55142 MW; 9F9770017C2455FC CRC64;

Query Match 55.1%; Score 54; DB 1; Length 507;
Best Local Similarity 60.0%; Pred. No. 0.53;
Matches 9; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 GHEQHQHGLGHGKXK 16
Db 145 GHGSHGHGHGHLAK 159

RESULT 13
ID ANTF SARPE STANDARD; PRT; 85 AA.
AC Q08617;

```

DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Antifungal protein precursor (AFP).
OS Sarcophaga peregrina (Flesh fly) (Boettcherisca peregrina).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Oestroidea;
OC Sarcophagidae; Sarcophaga.
OX NCBI_TaxID=7386;
RN [1]
RN SEQUENCE FROM N.A., AND SEQUENCE OF 20-54.
RC TISSUE=Fat body;
RX MEDLINE=93280179; PubMed=8505329;
RA Iijima R., Kurata S., Natori S.;
RT "Purification, characterization, and cDNA cloning of an antifungal
RT protein from the hemolymph of Sarcophaga peregrina (flesh fly)
RT larvae.";
RL J. Biol. Chem. 268:12055-12061(1993).
CC CC -1- FUNCTION: This protein inhibits the growth of a variety of
CC fungal species. The antifungal activity of this protein is
CC enhanced by the presence of sarcotoxin IA.
CC CC -1- SUBUNIT: Homodimer.
CC CC -1- TISSUE SPECIFICITY: Hemolymph.
CC CC -1- PTM: The N-terminus is blocked.
CC CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D13797; BAA02954.1; -.
DR PIR; A45969; A45969.
KW Fungicide; Signal; Repeat.
FT SIGNAL 1 18 POTENTIAL.
FT CHAIN 19 85 ANTIFUNGAL PROTEIN.
FT DOMAIN 19 73 2 X 7 AA REPEATS OF Q-H-G-H-G-G-Q.
FT REPEAT 19 25 1.
FT REPEAT 67 73 2.
SQ SEQUENCE 85 AA; 9018 MW; E381779F923F869B CRC64;

Query Match 54.1%; Score 53; DB 1; Length 85;
Best Local Similarity 57.1%; Pred. No. 0.12;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 HGHEQHQHGLGHGK 14
|||:|||||:
DB 22 HGGQDHGYGEGQ 35

RESULT 14
KE4 HUMAN
ID KE4 HUMAN STANDARD; PRT; 469 AA.
CD Q9UIQ0;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Zinc transporter SLC39A7 (Solute carrier family 39 member 1)
DE (Histidine-rich membrane protein Ke4).
GN SLC39A7 OR HKE4 OR RING5.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RN SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=97001166; PubMed=8812499;
RA Ando A., Kikui Y.Y., Shigenari A., Kawata H., Okamoto N., Shiina T.,
RA Chen L., Ikemura T., Abe K., Kimura M., Inoko H.;
RT "cDNA cloning of the human homologues of the mouse Ke4 and Ke6 genes

```

```

RT at the centromeric end of the human MHC region.";
RL Genomics 35:600-602(1996).
RN [2]
RP SEQUENCE FROM N.A.
RA Vergara A., Jana I., Corella A., de Miguel C., Migliaccio M.,
RA Encio I.;
RT "Molecular cloning and characterization of the human KE4 gene and 5'
RT flanking region.";
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RN SEQUENCE FROM N.A.
RA Tubby B.;
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE=Skin;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heieh F.,
RA Diatchenko L., Narusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Faney M., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [5]
RP CHARACTERIZATION.
RX PubMed=14525538;
RA Taylor K.M., Morgan H.B., Johnson A., Nicholson R.I.;
RT "Structure-function analysis of HKE4, a member of the new LIV-1
RT subfamily of zinc transporters.";
RL Biochem. J. 0:0-0(2003).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
CC -1- TISSUE SPECIFICITY: Major expression in placenta, lung, kidney
CC and pancreas.
CC -1- SIMILARITY: Belongs to the ZIP transporter (TC 2.A.5) family.
CC KE4/Catsup subfamily.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D82040; BAA11528.1; -.
CC EMBL; AF119221; AAD12305.1; -.
CC EMBL; AL031228; CAAC20238.1; -.
CC EMBL; BC000645; AAH00645.1; -.
CC Genew; HGNC:4927; SLC39A7.
CC MIM; 601416; -.
CC GO; GO:0005624; C:membrane fraction; TAS.
CC InterPro; IPR003689; Zn_transp_Zip.
CC Pfam; PF02535; Zip; 1.
KW Transport; Transmembrane; Glycoprotein.
FT TRANSMEM 10 30 POTENTIAL.
FT TRANSMEM 138 158 POTENTIAL.
FT TRANSMEM 169 189 POTENTIAL.
FT TRANSMEM 214 234 POTENTIAL.
FT TRANSMEM 381 401 POTENTIAL.

```

```

FT TRANSMEM 417 436 POTENTIAL.
FT DOMAIN 30 114 HIS-RICH.
FT DOMAIN 238 263 HIS-RICH.
FT CARBOHYD 330 330 N-LINKED (GLCNAC, ...) (POTENTIAL).
FT CONFLICT 7 7 A -> G (IN REF. 1 AND 2).
FT CONFLICT 280 280 E -> G (IN REF. 1 AND 2).
FT CONFLICT 376 376 S -> T (IN REF. 1 AND 2).
FT CONFLICT 397 469 CALLTEGAGVSGIAGGPGWLPFTAGGFIIVATVSVLP
FT ELRLASPLSLLEVLGLGVIMVLIABLE -> VPSSL
FT KEEQTVKLQVVQLAGSCHLLQVALST (IN REF. 1
FT AND 2).
FT SEQUENCE 469 AA; 50118 MW; 6504AJEF5AAG6A5B9 CRC64;
Query Match 54.1%; Score 53; DB 1; Length 469;
Best Local Similarity 57.1%; Pred. No. 0.69; Indels 0; Gaps 0;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 HGHEQHQHGLGHGX 14
Db 56 HGSHAHGHGTHE 69

RESULT 15
IR3 ARATH
ID IR3 ARATH STANDARD; PRT; 389 AA.
AC Q8LB59; Q8W244; Q9C957;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DE Fe(II) transport protein 3, chloroplast precursor (iron-regulated
transporter 3).
GN IR3 OR AT1G03960 OR T7P1.10.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP MEDLINE=21392307; PubMed=11500563;
RA Maeser P., Thomine S., Schroeder J.I., Ward J.M., Hirschi K., Sze H.,
RA Talke I.N., Amtmann A., Maathuis F.J.M., Sanders J.P.,
RA Tchieru J., Gribov M., Persans M.W., Salt D.E., Kim S.A.,
RA Gueriot M.L.;
RA "Phylogenetic relationships within cation transporter families of
Arabidopsis.";
RL Plant Physiol. 126:1646-1667(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=21016719; PubMed=11130712;
RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RA White O., Alonso J., Altafi H., Araujo R., Bowman C.L., Brooks S.Y.,
RA Buehler E., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
RA Chung M.K., Conn L., Conway A.B., Creasy T.H., Dewar K.,
RA Dunn P., Egu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,
RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
RA Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marzilli A.,
RA Miltschker J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
RA Pal G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
RA Utterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
thaliana.";
RL Nature 408:816-820(2000).
RN [3]
RP SEQUENCE FROM N.A.
RA Brover V., Troughan M., Alexandrov N., Lu Y.-P., Flavell R.,

```

```

RA Feldmann K.A.;
RT "Full-length cDNA from Arabidopsis thaliana.";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: May play a role in the transport of iron in the
CC plastids (By similarity).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Chloroplast
CC thylakoid membrane (POTENTIAL).
CC -!- SIMILARITY: Belongs to the ZIP transporter (TC 2.A.5) family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC EMBL; AF369915; AAL38498.1; ALT_INIT.
CC EMBL; AC018908; AAG51647.1; -.
CC EMBL; AY085604; AAM62825.1; -.
CC PIR; B96635; B96635.
CC InterPro; IPR004698; ZIP transport.
CC InterPro; IPR003689; Zn_transpt_Zip.
CC Pfam; PF02535; Zip; 1.
CC TIGRFAMs; TIGR00820; zip; 1.
CC Transport; Iron transport; Chloroplast; Transit peptide; Thylakoid;
CC Transmembrane.
CC TRANSIT 1 ? CHLOROPLAST (POTENTIAL).
CC CHAIN ? 389 FE(II) TRANSPORT PROTEIN 3.
CC DOMAIN ? 28 LUMENAL (POTENTIAL).
CC TRANSMEM 29 49 POTENTIAL.
CC DOMAIN 50 61 CYTOPLASMIC (POTENTIAL).
CC TRANSMEM 62 82 POTENTIAL.
CC DOMAIN 83 101 LUMENAL (POTENTIAL).
CC TRANSMEM 102 122 POTENTIAL.
CC DOMAIN 123 233 CYTOPLASMIC (POTENTIAL).
CC TRANSMEM 234 254 POTENTIAL.
CC DOMAIN 255 265 LUMENAL (POTENTIAL).
CC TRANSMEM 266 286 POTENTIAL.
CC DOMAIN 287 297 CYTOPLASMIC (POTENTIAL).
CC TRANSMEM 298 318 POTENTIAL.
CC DOMAIN 319 333 LUMENAL (POTENTIAL).
CC TRANSMEM 334 354 POTENTIAL.
CC DOMAIN 355 368 CYTOPLASMIC (POTENTIAL).
CC TRANSMEM 369 389 POTENTIAL.
CC SITE 247 248 HEAVY METALS BINDING (BY SIMILARITY).
CC CONFLICT 19 19 R -> L (IN REF. 3).
CC CONFLICT 77 77 G -> C (IN REF. 3).
CC SEQUENCE 389 AA; 41034 MW; 2CFF27202B5D996 CRC64;
Query Match 53.1%; Score 52; DB 1; Length 389;
Best Local Similarity 61.5%; Pred. No. 0.81;
Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 HGHEQHQHGLGHGX 13
Db 206 HARAHGHGHGHGX 218

Search completed: March 4, 2004, 13:08:06
Job time : 8.75 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model
Run on: March 4, 2004, 13:02:03 ; Search time 32 seconds
(without alignments)
167.619 Million cell updates/sec

Title: SEQ-A

Perfect score: 98

Sequence: 1 HGHEQHQHGLGHGKXKL 17

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: sp:archaea.*
- 2: sp:bacteria.*
- 3: sp:fungi.*
- 4: sp:human.*
- 5: sp:invertebrate.*
- 6: sp:mammal.*
- 7: sp:mhc.*
- 8: sp:organelle.*
- 9: sp:phase.*
- 10: sp:plant.*
- 11: sp:rodent.*
- 12: sp:virus.*
- 13: sp:vertebrate.*
- 14: sp:unclassified.*
- 15: sp:rvirus.*
- 16: sp:bacteriap.*
- 17: sp:archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	79	80.6	126	11 Q09016	Q09016 rattus norv
2	64	65.3	325	5 Q8MR03	Q8MR03 drosophila
3	62	63.3	314	16 Q86731	Q86731 streptomyces
4	62	63.3	361	16 Q9ABC7	Q9ABC7 caulobacter
5	60	61.2	79	10 Q9M435	Q9M435 quercus rob
6	60	61.2	574	16 Q828T3	Q828T3 streptomyces
7	60	61.2	911	5 Q86H89	Q86H89 dictyosteli
8	59	60.2	99	10 Q9ZRC7	Q9ZRC7 alnus eluti
9	59	60.2	349	5 Q9W2X1	Q9W2X1 drosophila
10	59	60.2	398	10 Q91036	Q91036 arabidopsis
11	59	60.2	583	5 Q9W4I6	Q9W4I6 drosophila
12	59	60.2	605	5 Q77280	Q77280 drosophila
13	59	60.2	1122	5 Q8SKD4	Q8SKD4 drosophila
14	58	59.2	199	10 Q9LYE2	Q9LYE2 arabidopsis
15	58	59.2	1064	5 Q9V5N1	Q9V5N1 drosophila
16	58	59.2	1085	5 Q24455	Q24455 drosophila

Q95p14 trypanosoma
Q99tn0 drosophila
Q9v3p9 drosophila
Q94dl7 oryza sativ
Q8pfc3 xanthomonas
Q9vwx5 drosophila
Q84nn0 oryza sativ
Q27920 bradyzia hy
Q9vui9 drosophila
Q9w4c1 drosophila
Q9vws0 drosophila
Q9nes7 caenorhabdi
Q8irn6 drosophila
Q9w3l3 drosophila
Q9cqc6 mus musculu
Q9d6h6 mus musculu
Q9xel3 picea glauc
Q9p112 campylobact
Q9v3d3 talstonia s
Q8tch2 homo sapien
Q8new0 homo sapien
Q9jkn1 mus musculu
Q80y27 mus musculu
Q84l19 oryza sativ
Q947r8 eucalyptus
Q26227 rhynchosia
Q96w52 emeritella
Q9vyp3 drosophila
Q9nnv9 plasmodium

ALIGNMENTS

RESULT 1
ID Q09016 PRELIMINARY; PRT: 126 AA.
AC Q09016;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE K-kininogen (Fragment).
GN KNGK.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Donryu;
RX MEDLINE=97468288; PubMed=9321484;
RA Harris E.L., Grigor M.R., Innes B.A., Harrop S.B., Koike G.,
RA Jacob H.J.;
RT "Strain-specific deletions in exon 10 of rat K-kininogen and T1-
RT kininogen genes allow mapping of both genes to rat chromosome 11.";
RL Mamm. Genome 8:791-792(1997).
DR EMBL; AF003623; AAC09070.1; -.
DR InterPro; IPR002395; Kininogen.
DR PRINTS; PR00334; KININOGEN.
DR NON TER 1 1
FT VARIANT 24 24 D -> N.
FT VARIANT 87 89 RDK -> HQG.
FT VARIANT 115 115 S -> N.
FT NON TER 126 126
SQ SEQUENCE 126 AA; 9CCDF8751DA49C88 CRC64;
Query Match 80.6%; Score 79; DB 11; Length 126;
Best Local Similarity 70.8%; Pred. No. 4.6e-05;
Matches 12; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 HGHEQHQHGLGHGKXKL 17

Db 38 HGHEQHQHGLGHGKXKL 54

RESULT 2

Q8MR03
ID Q8MR03 PRELIMINARY; PRT; 325 AA.
AC Q8MR03;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE LP03275p (Fragment).
GN CAUP OR CG10605.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RA Stapleton M., Brokstein P., Hong L., Achbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Drensek D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nuno J., Pacleb J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Ceiniker S.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY12206; AAM52718.1; -.
DR FlyBase; FBgn0015919; caup.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR003893; Iroquois_homeo.
DR SMART; SM00548; IRO; 1.
FT NON TER 1
SQ SEQUENCE 325 AA; 3504 MW; 07AC5FF3D4322FCD CRC64;

Query Match 65.3%; Score 64; DB 5; Length 325;
Best Local Similarity 76.9%; Pred. No. 0.032;
Matches 10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 HGHGQHQHGLGHGH 13

||| |||||
Db 288 HGHGQHQHGLGHGH 300

RESULT 3

O86731
ID O86731 PRELIMINARY; PRT; 314 AA.
AC O86731;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein SC06549.
GN SC06549 OR SC5C7.34.
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Seeger K.J., Harris D.;
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Parkhill J., Barrell B.G., Rajandream M.A.;
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RX MEDLINE=97000351; PubMed=8843436;
RA Redenbach M., Kieser H.N., Denapate D., Eichner A., Cullum J.,

RA Kinashi H., Hopwood D.A.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL Mol. Microbiol. 21:77-95(1996).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RX MEDLINE=21996410; PubMed=12000953;
RA Bentley S.D., Chater K.F., Cerrido-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Murphy L., Oliver K., O'Neil S.,
RA Huang C.-H., Kieser T., Lark L., Lark L., Lark L., Lark L., Lark L.,
RA Rabinowitch E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2).";
RL Nature 417:141-147(2002).
DR EMBL; AL939128; CAA20646.1; -.
DR PIR; T35241; T35241.
DR GO; GO:0008757; F:S-adenosylmethionine-dependent methyltransf.; IEA.
DR InterPro; IPR000051; SAM_bind.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 314 AA; 33586 MW; 7CA3288CC28FF007 CRC64;

Query Match 63.3%; Score 62; DB 16; Length 314;
Best Local Similarity 76.9%; Pred. No. 0.065;
Matches 10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 HGHGQHQHGLGHGH 13

||| |||||
Db 18 HGHGQHQHGLGHGH 30

RESULT 4

Q9ABC7
ID Q9ABC7 PRELIMINARY; PRT; 361 AA.
AC Q9ABC7;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Cation efflux family protein.
GN CC0303.
OS Caulobacter crescentus.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacterales;
OC Caulobacteraceae; Caulobacter.
OX NCBI_TaxID=155892;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 19089 / CB15;
RX MEDLINE=21173698; PubMed=11259647;
RA Nierman W.C., Feldblyum T.V., Laub M.R.K., Ohta N., Maddock J.R.,
RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
RA Deboy R.T., Dodson R.J., Durkin A.S., Ghim M.L., Haft D.H.,
RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,
RA Tutterback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,
RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
RT "Complete genome sequence of Caulobacter crescentus";
RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
DR EMBL; AE005704; AAK2290.1; -.
DR PIR; F87286; F87286.
DR TIGR; CC0303; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0008224; P:cation transporter activity; IEA.
DR GO; GO:0006812; P:cation transport; IEA.
DR InterPro; IPR002524; Cation_efflux.
DR Pfam; PF01545; Cation_efflux; 1.
DR TIGRFAMs; TIGR01297; CDF; 1.
KW Complete proteome.
SQ SEQUENCE 361 AA; 38180 MW; 1A4F7F0A7C62EBE0 CRC64;

```

Query Match      63.3%; Score 62; DB 16; Length 361;
Best Local Similarity 69.2%; Pred. No. 0.076; 3; Indels 0; Gaps 0;
Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 HGHEQHQHGLGHGH 13
DB 39 HGHDAHGHGHGH 51

RESULT 5
Q9M435 PRELIMINARY; PRT; 79 AA.
ID Q9M435
AC Q9M435;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Phase-change related protein precursor.
OS Quercus robur (English oak).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fagales; Fagaceae; Quercus.
OX NCBI_TaxID=38942;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=In vitro shoot cultures;
RA Gil B.; Pastoriza E.M.; Ballester A.; Sanchez C.; oak shoot cultures
RT "Identification of a phase-change related mRNA in oak shoot cultures
RT derived from basal sprouts and crown branches."
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ271778; CAB72442.1; -.
KW Signal.
FT SIGNAL.
SQ SEQUENCE 79 AA; 8414 MW; 8B45CABF40F0B6F CRC64;

Query Match      61.2%; Score 60; DB 10; Length 79;
Best Local Similarity 62.5%; Pred. No. 0.03; 6; Indels 0; Gaps 0;
Matches 10; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 HGHEQHQHGLGHGHKK 16
DB 48 HGHGHGHGHGHGHGHGK 63

RESULT 6
Q828T3 PRELIMINARY; PRT; 574 AA.
ID Q828T3
AC Q828T3;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
DE Hypothetical protein.
GN SAV6579.
OS Streptomyces avermitilis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycinae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=33903;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=21477403; PubMed=11572948;
RA Omura S.; Ikeda H.; Ishikawa J.; Hanamoto A.; Takahashi C.;
RA Shinose M.; Takahashi Y.; Horikawa H.; Nakazawa H.; Osonoe T.;
RA Kikuchi H.; Shiba T.; Sakaki Y.; Hattori M.;
RT "Genome sequence of an industrial microorganism Streptomyces
RT avermitilis: deducing the ability of producing secondary
RT metabolites."
RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=NA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=22608306; PubMed=12692562;
RA Ikeda H.; Ishikawa J.; Hanamoto A.; Shinose M.; Kikuchi H.; Shiba T.;

```

```

Sakaki Y., Hattori M., Omura S.;
"Complete genome sequence and comparative analysis of the industrial
microorganism Streptomyces avermitilis."
Nat. Biotechnol. 21:526-531(2003).
EMBL; AP005047; BAC74290.1; -.
GO; GO:0016021; C:integral to membrane; IEA.
GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
GO; GO:0007186; P:G-protein coupled receptor protein signalin. .; IEA.
InterPro; IPR000276; GPCR_Rhodopsn.
InterPro; IPR000437; Prok_lipoprot_S.
PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
PROSITE; PS00013; PROKAR_LIPOPROTEIN_1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 574 AA; 59898 MW; 8061208EG33EE9E CRC64;

Query Match      61.2%; Score 60; DB 16; Length 574;
Best Local Similarity 64.3%; Pred. No. 0.26; 4; Indels 0; Gaps 0;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 HGHEQHQHGLGHGHK 14
DB 418 HKHNENHGHGHGHK 431

RESULT 7
Q86H89 PRELIMINARY; PRT; 911 AA.
ID Q86H89
AC Q86H89;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
DE Hypothetical protein.
OS Dictyostelium discoideum (slime mold).
OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AX4;
RX MEDLINE=22092622; PubMed=12097910;
RA Gloeckner G.; Eichinger L.; Szafranski K.; Pachebat J.; Dear P.;
RA Lehmann R.; Baumgart C.; Parra G.; April J.F.; Guigo R.; Kumpf K.;
RA Tungal B.; Cox E.; Quail M.A.; Platzer M.; Rosenthal A.; Noegel A.A.;
RT "Sequence and analysis of chromosome 2 of Dictyostelium discoideum."
RL Nature 418:79-85(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=AX4;
RA Baumgart C.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC116960; AA053134.1; -.
DR InterPro; IPR007087; Znf_C2H2.
DR Pfam; PF00096; zf-C2H2; 2.
DR SMART; SM00355; Znf_C2H2; 4.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 2.
KW Hypothetical protein.
SQ SEQUENCE 911 AA; 102289 MW; DA816C7DB6435B16 CRC64;

Query Match      61.2%; Score 60; DB 5; Length 911;
Best Local Similarity 69.2%; Pred. No. 0.44; 4; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 HGHEQHQHGLGHGH 13
DB 133 HGHHAHGHGHGH 145

RESULT 8
Q9ZRC7 PRELIMINARY; PRT; 99 AA.
ID Q9ZRC7
AC Q9ZRC7;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

```


DE Actinorhizal nodulin AgNOD-GHRP.
GN AGN184.
OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fagales; Betulaceae; Alnus.
OX NCBI_TaxID=3517;
[1]
RN SEQUENCE FROM N.A.
RP TISSUE=Root nodules;
RC Dobritsa S.V., Mullin B.C.;
RT "In vitro expression of actinorhizal nodulin AgNOD-GHRP and
RT demonstration of its toxicity to *Escherichia coli*.";
RL (in) Stacey G., Mullin B.C., Grasshoff P.M. (eds.);
RL THE BIOLOGY OF PLANT-MICROBE INTERACTIONS:
RL PROCEEDINGS OF THE 8TH INTERNATIONAL SYMPOSIUM ON MOLECULAR
RL PLANT-MICROBE INTERACTIONS, pp.1-1. Unknown Publisher (1996).
[2]
RN SEQUENCE FROM N.A.
RP TISSUE=Root nodules;
RC Dobritsa S.V., Mullin B.C.;
RT "Isolation of a nodule-specific cDNA encoding a putative glycine-rich
RT protein from *Alnus glutinosa*.";
RL Thesis (1993), The University of Tennessee, Knoxville, TN, USA.
[3]
RN SEQUENCE FROM N.A.
RP TISSUE=Root nodules;
RC Pawlowski K., Twigg P.G., Dobritsa S.V., Guan C., Mullin B.C.;
RT "A nodule-specific gene family from *Alnus glutinosa* encodes glycine
RT and histidine-rich proteins expressed in the early stages of
RT actinorhizal nodule development.";
RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U69156; AAC00171.1; -.
SQ SEQUENCE 99 AA; 10567 MW; 2ACBE4D57C070E83 CRC64;
Query Match 60.2%; Score 59; DB 10; Length 99;
Best Local Similarity 69.2%; Pred. No. 0.056;
Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
Qy 1 HGHEQVQGLGHGH 13
||| ||| |||
Db 50 HGHEVHVGHGHGH 62
RESULT 9
ID Q9W2X1 PRELIMINARY; PRT; 349 AA.
AC Q9W2X1; Q96082;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE CG2961 protein (UD38046p).
GN CG2961.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=Berkley;
RC Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter B.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bertram B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,

Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
de Pablos B., Delcher A., Deng I., Mays A.D., Dew I., Dietz S.M.,
Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Durkin B.C., Dunn P.,
Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
Harris N.D., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai X.,
Lasko P., Lei Y., Levitsky A.C., Li J., Li Z., Liang Y., Lin X.,
Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson B.,
Merkulov G., Mlsina N.V., Mobarry C., Morris J., Moshrefi A.,
Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
Nelson D.R., Nelson K.A., Nixon K., Ruskern D.R., Pacleb J.M.M.G.,
Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese H.,
Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
Spieler E., Spradling A.C., Stapleton M., Strong R., Sun E.,
Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J., Yao Q.A.,
Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of *Drosophila melanogaster*.";
Science 287:2185-2195 (2000).
[2]
RN SEQUENCE FROM N.A.
RP Celniker S.E., Adams M.D., Kroumiller B., Wan K.H., Holt R.A.,
RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
RA Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,
RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
RA Dodson K., Dorsett V., Dou P.L.E., Doyle C., Dresnek D., Farfan D.,
RA Ferrera S., Frise E., Galle R.F., Garg N.S., George R.A.,
RA Gonzalez M., Houch J., Hoskins R.A., Hostin D., Howland T.J.,
RA Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunco J.,
RA Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B., Scheeler F.,
RA Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
RA Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,
RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
RT "Sequencing of *Drosophila melanogaster* genome.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
[3]
RN SEQUENCE FROM N.A.
RP Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
RA Hradecky P., Huang Y., Kaminker J.S., Prochuk S.E., Smith C.D.,
RA Tuzy J.L., Bergman C., Berman B., Carlson J.W., Celniker S.E.,
RA Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
RA Kronmiller B., Marshall B., Millburn G., Richter J., Russo S.,
RA Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,
RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
RT "Annotation of *Drosophila melanogaster* genome.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
[4]
RN SEQUENCE FROM N.A.
RP Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
[5]
RN SEQUENCE FROM N.A.
RP FlyBase;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
[6]
RN SEQUENCE FROM N.A.
RP STRAIN=Berkley;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA Nuno J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,
RA Yu C., Lewis S.E., Rubin G.M., Celniker S.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
EMBL; AE003449; AAF46566.2; -.

QY	1	HGHEQCHLGHGH	13	
Db	184	HGHSHGHGHGHGH	196	

RESULT 11				
Q9W416				
AC	Q9W416	PRELIMINARY;	PRT;	583 AA.
ID	Q9W416;			
DT	01-MAY-2000	(TRENBLrel. 13, Created)		
DT	01-OCT-2002	(TRENBLrel. 22, Last sequence update)		
DT	01-MAR-2003	(TRENBLrel. 23, Last annotation update)		
DE	CG3081 protein.			
GN	EG:84H4. 4 OR CG3081.			
OS	Drosophila melanogaster (Fruit fly).			
OC	Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;			
OC	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;			
OC	Ephydroidea; Drosophilidae; Drosophila.			
OX	NCBI_TaxID=7227;			
EN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=Berkeley;			
RC	MEDLINE=20196006; PubMed=10731132;			
RX	Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,			
RA	Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Gallie R.F.,			
RA	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,			
RA	Sutton G.G., Wortman J.R., Vandell M.D., Zhang Q., Chen L.X.,			
RA	Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,			
RA	Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,			
RA	Abiril J.F., Agbayani A., An H.-J., Andrews-Ffankoch C., Baldwin D.,			
RA	Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,			
RA	Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,			
RA	Borkova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,			
RA	Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,			
RA	Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,			
RA	de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,			
RA	Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,			
RA	Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,			
RA	Foeller C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,			
RA	Glocke A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,			
RA	Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,			
RA	Hoslin D., Houshou K.A., Howland T.J., Wei M.-H., Ibegwam C.,			
RA	Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,			
RA	Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,			
RA	Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,			
RA	Liu X., Mactel B., McIntosh T.C., McLeod M.P., McPherson D.,			
RA	Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,			
RA	Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,			
RA	Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacieb J.M.,			
RA	Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,			
RA	Reinert K., Remington K., Saunders R.D.C., Scheeler P., Shen H.,			
RA	Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,			
RA	Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,			
RA	Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,			
RA	Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,			
RA	Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,			
RA	Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,			
RA	Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,			
RA	Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;			
RT	"The genome sequence of Drosophila melanogaster.";			
RT	Science 287:2185-2195(2000).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RA	Celisner S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,			
RA	Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,			
RA	Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,			
RA	Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,			
RA	Dodson K., Dorsett V., Doup L.E., Doyle C., Dresek D., Farfan D.,			
RA	Ferreira S., Frise E., Gallie R.F., Garg N.S., George R.A.,			
RA	Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,			
RA	Ibegwam C., Jalali M., Kruse D., Li P., Mactel B., Moshrefi A.,			
RA	McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunco J.,			

```

RA Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B.,
RA Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
RA Stapleton M., Strong R., Swirskas R., Tector C., Tyler D.,
RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
RT "Sequencing of Drosophila melanogaster genome."
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
RA Hradecky P., Huang Y., Kamman J.S., Prochuk S.E., Smith C.D.,
RA Tupy J.L., Bergman C., Rankin B., Carlson J.W., Celnik S.E.,
RA Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
RA Krommiller B., Marshall B., Millburn G., Richter J., Russo S.,
RA Searle S.M.J., Smith E., Snu S., Smutniak F., Whitfield E.,
RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
RT "Annotation of Drosophila melanogaster genome."
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Adams M.D., Celnik S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RA FlyBase;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF003431; AAF45965.2; -.
DR FlyBase: FBgn0025613; EG:84H4.4.
SQ SEQUENCE 583 AA; 62897 MW; 0B57A165245C2D5C CRC64;

Query Match 60.2%; Score 59; DB 5; Length 583;
Best Local Similarity 69.2%; Pred. No. 0.39;
Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 HGHQHQHGHGHGH 13
Db 467 HGHGHGHGHGHGH 479

RESULT 12
O77280 PRELIMINARY; PRT; 605 AA.
AC O77280;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE EG:84H4.4 protein.
GN EG:84H4.4 OR CG3081.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA Ferraz C., Vidal S., Brun C., Bucheton A., Demaille J.G.;
RT "Sequencing the discal X chromosome of Drosophila melanogaster."
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Beros P.;
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL031766; CAA21135.1; -.
DR FlyBase: FBgn0025613; EG:84H4.4.
SQ SEQUENCE 605 AA; 64947 MW; B06C84AC4D7D2C64 CRC64;

Query Match 60.2%; Score 59; DB 5; Length 605;
Best Local Similarity 69.2%; Pred. No. 0.4;
Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 HGHQHQHGHGHGH 13
Db 341 HGHGHGHGHGHGH 353

RESULT 13
O85XD4 PRELIMINARY; PRT; 1122 AA.
AC O85XD4; Q9W287;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE GH03394P (CG32676 protein).
GN CG32676 OR CG9725 OR CG9732.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RA Stapleton M., Brokstein P., Hong L., Agbavani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan X., Yu C., Lewis S.E., Rubin G.M.,
RA Celnik S.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazew R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbavani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
RA Balaw R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrera S., Fleisschmann W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,
RA Jatali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Riden-Kiamos I., Simpson M., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Spiers R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Yeh J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
RN [3]
RP SEQUENCE FROM N.A.
RA Celnik S.E., Adams M.D., Krommiller B., Wan K.H., Holt R.A.,
RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
RA Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busan D.A.,
RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,

```

GO: GO:0005489; P:electron transporter activity; IEA.
GO: GO:0006118; P:electron transport, IEA.
GO: InterPro; IPR000345; CytC heme_BS.
PROSITE; PS00190; CYTOCHROME_C; 1.
KW Hypothetical protein.
SQ SEQUENCE 199 AA; 21539 MW; E5D28AC167B3FBF8 CRC64;

Query Match 59.28; Score 58; DB 10; Length 199;
Best Local Similarity 69.28; Pred. No. 0.17; 3; Indels 0; Gaps 0;
Matches 9; Conservative 1; Mismatches 0;

1 HGHFQQHGLGHHG 13
||| : |||||
92 HGHGRGHGHGHG 104

RESULT 15

Q9V5N1 PRELIMINARY; PRT; 1064 AA.
ID Q9V5N1 Q9V5N2; Q24194; Q24187;
AC Q9V5N1; Q9V5N2; Q24194; Q24187;
CD 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Psg protein (F033470P).
DE PSQ OR CG2368.
OS Drosophila melanogaster (Fruit fly).
OS Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
ON NCBI_TaxID=7227;
RX [1]
RN SEQUENCE FROM N.A. (ISOFORM A).
RC TISSUE=ovary;
RP MELINE=95220671; PubMed=7705633;
RX Horowitz H., Berg C.A.;
RT "Aberrant splicing and transcription termination caused by P element
RT insertion into the intron of a drosophila gene.";
RL Genetics 139:327-335(1995).
RN [2]
RN SEQUENCE FROM N.A. (ISOFORMS A AND B).
RC TISSUE=ovary;
RP MELINE=96232300; PubMed=8674425;
RX Horowitz H., Berg C.A.;
RT "The Drosophila pipsqueak gene encodes a nuclear BTB-domain-containing
RT protein required early in oogenesis.";
RL Development 122:1859-1871(1996).
RN [3]
RN SEQUENCE FROM N.A. (ISOFORMS A AND 2).
RC STRAIN=Berkeley; PubMed=10731132;
RX MELINE=20196006; PubMed=10731132;
RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.P.,
RA George R.A., Lewis S.B., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.P., Adayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Benson K.Y., Bencos P.V., Bereman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Paulos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA Foaier C., Gabrielian A.B., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hootin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,
RA Jajali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.C., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,

RA Merklov G., Milshina N.V., Moharry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Murry D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M., M.G.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-E., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhou W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkeley;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA Nunoo J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,
RA Yu C., Lewis S.E., Rubin G.M., Celniker S.;
RL Submitted (DEC-2001) to the EMBL/Genbank/DBJ databases.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=3;
CC Name=H; Synchrony=1;
CC Name=H; Synchrony=1; Sequence=Displayed;
CC IsoId=Q9V5N1-1; Sequence=VSP_050295;
CC Name=H; Synchrony=1;
CC IsoId=Q9V5N1-2; Sequence=VSP_050296;
CC Name=H; Synchrony=1;
CC IsoId=Q9V5N1-3; Sequence=VSP_050296;
CC EMBL; U48358; AAC47153.1; -;
CC EMBL; U48402; AAC47154.1; -;
CC EMBL; AE003829; AAF58769.1; -;
CC EMBL; AE003829; AAF58770.1; -;
CC EMBL; AY069588; AAL39733.1; -;
CC FlyBase; FBgn004399; psq.
CC GO; GO:000515; P:protein binding; IEA.
CC GO; GO:0003700; P:transcription factor activity; IEA.
CC GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
CC InterPro; IPR00210; BTB_POZ.
CC InterPro; IPR002197; HTH_Fis.
CC InterPro; IPR007889; HTH_psq.
CC Pfam; PF00653; BTB; 1.
CC Pfam; PF05225; HTH_Dsq; 4.
CC SMART; SM00225; BTB; 1.
CC TIGRFAMs; TIGR01199; HTH_fis; 2.
CC PROSITE; PS50097; BTB; 1.
KW Alternative splicing.
FT VARSPLIC 1 429 Missing (in isoform B).
FT VARSPLIC 719 736 /FTid=VSP_050295.
FT VARSPLIC 719 736 Missing (in isoform 2).
FT CONFLICT 1020 1020 /FTid=VSP_050296.
FT CONFLICT 1020 1020 Q -> QQ (IN REF. 1 AND 2).
SQ SEQUENCE 1064 AA; 114984 MW; 77420C782DE6CA5 CRC64;

Query Match 59.2%; Score 58; DB 5; Length 1064;
Best Local Similarity 69.2%; Pred. No. 1.1;
Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 HGHEQHQGLGHGH 13
DB 332 HEHEHGHGHGH 344

Search completed: March 4, 2004, 13:10:27
Job time : 34 secs

Result No.	Query		DB	ID	Description
	Score	Match Length			
1	160	98.8	28	3	Aay81996 Human hig
2	160	98.8	55	3	Aay93346 Light cha
3	160	98.8	62	3	Aay93348 Light cha
4	160	98.8	63	2	Aay75186 Partial p
5	160	98.8	83	3	Aay93347 Light cha
6	160	98.8	94	3	Aay93351 Light cha
7	160	98.8	125	5	Abb78708 Human hig
8	160	98.8	131	2	Aay75181 Partial p
9	160	98.8	158	5	Abb78709 Calmoduli
10	160	98.8	186	3	Aay93349 Light cha
11	160	98.8	255	3	Aay93342 Light cha
12	160	98.8	255	4	Aab73620 Human hig
13	160	98.8	305	4	Abg21100 Novel hum
14	160	98.8	357	6	Abx41202 Human DIT
15	160	98.8	415	6	Abu99146 Novel hum
16	160	98.8	615	6	Abu99144 Novel hum
17	160	98.8	626	5	Abb78707 Human hig
18	160	98.8	644	4	Abg21101 Novel hum
19	160	98.8	644	5	Abb78710 Human hig
20	160	98.8	644	6	Abu99150 Novel hum
21	160	98.8	644	6	Abu99145 Novel hum
22	160	98.8	720	4	Abg21103 Novel hum
23	138	85.2	179	3	Aay93353 Light cha
24	91	56.2	47	3	Aay93345 Light cha
25	88	54.3	17	2	Aaw627 Human hig

```

SQ      Sequence 55 AA;
Query Match      98.8%;      Score 160;      DB 3;      Length 55;
Best Local Similarity 96.4%;      Prod. No. 1.7e-15;
Matches 27;      Conservative 0;      Mismatches 1;      Indels 0;      Gaps 0

```

RESULT 4
AAR75186
ID AAR75186 standard; peptide; 63 AA.
XX
XX AAR75186.

XX 05-DEC-1995 (first entry)
 XX Partial peptide of human HMW kininogen fragment 2.
 DE high molecular weight; kininogen; fragment; 1.2; 1; 2; partial;
 KW wound treating agent; bovine; growth promotion; fibroblast.
 XX
 OS Homo sapiens.
 XX JP07082172-A.
 PN 28-MAR-1995.
 XX 17-SEP-1993; 93JP-00230616.
 XX 17-SEP-1993; 93JP-00230616.
 XX (FARH) HOECHST JAPAN KK.
 XX WPI; 1995-156909/21.
 XX A wound treating agent contg. a partial peptide of kininogen - have
 PT growth promotion activity of fibroblasts.
 XX
 PS Claim 8; Page 8; 8pp; Japanese.
 XX AAR75186 is a partial peptide corresponding to human kininogen fragment
 CC 1, amino acids 458-520. Partial peptides of bovine and human kininogen
 CC fragments 1.2, 1 and 2, are used in wound treating agent compens. and act
 CC as the active component. The fragments are useful in wound treating
 CC because they have growth promotion activity on fibroblasts
 XX
 SQ Sequence 63 AA;
 Query Match 98.8%; Score 160; DB 2; Length 63;
 Best Local Similarity 96.4%; Pred. No. 2e-15;
 Matches 27; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 HGHEQQHGLGHGKXKLDLLEHQQGHV 28
 DB 6 HGHEQQHGLGHGKFLDDLLEHQQGHV 33
 RESULT 5
 AAY93347
 ID AAY93347 standard; peptide; 83 AA.
 AC AAY93347;
 XX
 DT 04-SEP-2000 (first entry)
 DE Light chain of human high molecular weight kininogen analogue.
 XX Human; high molecular weight kininogen; glycoprotein; endothelial cell;
 KW plasma kallikrein; heavy chain; light chain; analogue; angiogenesis;
 KW endothelial cell proliferation; endothelial cell migration; vitronectin.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX WO200027415-A2.
 XX 18-MAY-2000.
 XX 09-NOV-1999; 99WO-US026377.
 XX 10-NOV-1998; 98US-0107844P.
 XX (UTEM) UNIV TEMPLE.
 PA (DUPO) DUPONT PHARM CO.
 PA (COLM/) COLMAN W R.
 PA (MOUS/) MOUSA A S.
 XX
 PI Colman WR, Mousa AS;
 XX WPI; 2000-376306/32.
 XX Method for inhibiting endothelial cell proliferation, using compound that
 PT inhibit endothelial cell migration.
 XX Claim 8; Page 39; 41pp; English.
 XX

XX Colman WR, Mousa AS;
 XX WPI; 2000-376306/32.
 XX Method for inhibiting endothelial cell proliferation, using compound that
 PT inhibit endothelial cell migration.
 XX
 PS Claim 5; Page 37; 41pp; English.
 XX The present sequence represents an analogue of the light chain of human
 CC high molecular weight kininogen. High molecular weight kininogen is a 120
 CC kDa glycoprotein which binds with high affinity to endothelial cells,
 CC where it is cleaved by plasma kallikrein into heavy and light chains.
 CC Analogues of high molecular weight kininogen are used in the method of
 CC the invention. The specification describes a method of inhibiting
 CC endothelial cell proliferation. The method comprises contacting
 CC endothelial cells with a compound containing high molecular weight
 CC kininogen analogues. The method and the compounds can be used for
 CC inhibiting endothelial cell proliferation. The compounds can also be used
 CC for inhibiting angiogenesis. The compounds can also be used to inhibit
 CC migration of endothelial cells to vitronectin
 XX
 SQ Sequence 83 AA;
 Query Match 98.8%; Score 160; DB 3; Length 83;
 Best Local Similarity 96.4%; Pred. No. 2.7e-15;
 Matches 27; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 HGHEQQHGLGHGKXKLDLLEHQQGHV 28
 DB 26 HGHEQQHGLGHGKFLDDLLEHQQGHV 53
 RESULT 6
 AAY93351
 ID AAY93351 standard; peptide; 94 AA.
 XX
 AC AAY93351;
 XX
 DT 04-SEP-2000 (first entry)
 DE Light chain of human high molecular weight kininogen analogue.
 XX Human; high molecular weight kininogen; glycoprotein; endothelial cell;
 KW plasma kallikrein; heavy chain; light chain; analogue; angiogenesis;
 KW endothelial cell proliferation; endothelial cell migration; vitronectin.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX WO200027415-A2.
 XX 18-MAY-2000.
 XX 09-NOV-1999; 99WO-US026377.
 XX 10-NOV-1998; 98US-0107844P.
 XX (UTEM) UNIV TEMPLE.
 PA (DUPO) DUPONT PHARM CO.
 PA (COLM/) COLMAN W R.
 PA (MOUS/) MOUSA A S.
 XX
 PI Colman WR, Mousa AS;
 XX WPI; 2000-376306/32.
 XX Method for inhibiting endothelial cell proliferation, using compound that
 PT inhibit endothelial cell migration.
 XX Claim 8; Page 39; 41pp; English.
 XX

[illegible]

74 HGHEOOHGLGHGHKFKLDDLEHQGGHV 101

RESULT 9
 ABB78709
 ID ABB78709 standard; protein; 158 AA.
 AC ABB78709;
 XX 18-JUL-2002 (first entry)
 DE Calmodulin binding protein and D5 domain fusion protein SEQ ID NO:13.
 DE Human; kininogen; high molecular weight kininogen; HK; D5 domain;
 KW D5 receptor; angiotensin; endothelial cell; cytotatic; antitumour;
 KW antithrombotic; vasotrophic; vulnerary; tranquilliser; thrombolytic;
 KW ophthalmological; gynaecological; antitumor; antidiabetic; antiarthritic;
 KW antiangiogenic; antiapoptotic; endocrine; apoptosis; gene therapy;
 KW calmodulin binding protein; CBP.
 OS Homo sapiens.
 XX Key Location/Qualifiers
 FH Domain 1..29
 FT Domain /notes="Calmodulin binding protein (CBP) sequence"
 FT Domain 34..158
 FT Domain /label=D5_domain
 XX WO200214369-A2.
 XX 21-FEB-2002.
 XX 24-JUL-2001; 2001WO-US023185.
 XX 24-JUL-2000; 2000US-0220194P.
 XX (ATTE-) ATTENUON LLC.
 PA Mazar AP, Juarez JC;
 PI WPI; 2002-393611/42.
 DR Novel human kininogen D5 domain polypeptides useful for treating
 PT conditions associated with endothelial cell migration, proliferation,
 PT invasion or angiogenesis, e.g. arthritis, macular degeneration, benign
 PT hyperplasia.
 XX Example 3; Page 62; 84pp; English.
 CC The present invention describes an isolated polypeptide (I) that
 CC corresponds to the D5 domain of human kininogen, or biologically active
 CC peptide fragment, homologue or functional derivative, and which: (a)
 CC inhibits angiogenesis; (b) binds to the D5 binding site on endothelial
 CC cells (EC); (c) activates signalling pathways leading to the introduction
 CC of apoptosis in EC; and/or (d) inhibits the signalling pathway required
 CC for maintenance of EC viability. (I) has cytostatic, antitumour,
 CC antithrombotic, vasotrophic, vulnerary, tranquiliser, thrombolytic,
 CC ophthalmological, gynaecological, antitumor, antidiabetic, antiarthritic,
 CC antiangiogenic, antiapoptotic and endocrine activities. An antibody (IX)
 CC specific for an epitope of (I) is useful for inhibiting tumour growth or
 CC angiogenesis in a subject. (II), a D5 fusion polypeptide (II) or a dimeric
 CC or trimeric fusion polypeptide (III) can be used for inhibiting EC
 CC migration, proliferation, invasion, or angiogenesis, or for inducing EC
 CC apoptosis. An angiogenic EC-targeting pharmaceutical composition (X)
 CC comprising (I), (II), or (III), can be used for treating a subject having
 CC a disease or condition associated with undesired EC migration,
 CC proliferation, invasion or angiogenesis. (I), (II), or (III) can be used
 CC for isolating a D5 domain binding molecule from a complex mixture and for
 CC isolating or enriching cells expressing D5 domain binding sites from a
 CC cell mixture. The present sequence represents a calmodulin binding
 CC protein and D5 domain fusion protein, which is given in an example from
 CC the present invention
 XX Sequence 158 AA;

Query Match 98.8%; Score 160; DB 5; Length 158;
 Best Local Similarity 96.4%; Pred. No. 5.4e-15;
 Matches 27; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 HGHEQHQHGLGHGKXKLDLLEHQGHV 28
 DB 95 HGHEQHQHGLGHGKXKLDLLEHQGHV 122
 RESULT 10
 AAY93349
 ID AAY93349 standard; peptide; 186 AA.
 AC AAY93349;
 XX 04-SEP-2000 (first entry)
 DE Light chain of human high molecular weight kininogen analogue.
 DE Human; high molecular weight kininogen; glycoprotein; endothelial cell;
 KW plasma kallikrein; heavy chain; light chain; analogue; angiogenesis;
 KW endothelial cell proliferation; endothelial cell migration; vitronectin.
 OS Synthetic.
 OS Homo sapiens.
 PN WO200027415-A2.
 XX 18-MAY-2000.
 XX 09-NOV-1999; 99WO-US026377.
 XX 10-NOV-1998; 98US-0107844P.
 XX (UTEM) UNIV TEMPLE.
 PA (DUPO) DUPONT PHARM CO.
 PA (COLM) COLMAN W R.
 PA (MOUS) MOUSA A S.
 XX Colman WR, Mousa AS;
 XX WPI; 2000-376306/32.
 XX Method for inhibiting endothelial cell proliferation, using compound that
 XX inhibit endothelial cell migration.
 XX Claim 9; Page 38; 41pp; English.
 CC The present sequence represents an analogue of the light chain of human
 CC high molecular weight kininogen. High molecular weight kininogen is a 120
 CC kDa glycoprotein which binds with high affinity to endothelial cells,
 CC where it is cleaved by plasma kallikrein into heavy and light chains.
 CC Analogues of high molecular weight kininogen are used in the method of
 CC the invention. The specification describes a method of inhibiting
 CC endothelial cell proliferation. The method comprises contacting
 CC endothelial cells with a compound containing high molecular weight
 CC kininogen analogues. The method and the compounds can be used for
 CC inhibiting endothelial cell proliferation. The compounds can also be used
 CC for inhibiting angiogenesis. The compounds can also be used to inhibit
 CC migration of endothelial cells to vitronectin
 XX Sequence 186 AA;
 Query Match 98.8%; Score 160; DB 3; Length 186;
 Best Local Similarity 96.4%; Pred. No. 6.4e-15;
 Matches 27; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 HGHEQHQHGLGHGKXKLDLLEHQGHV 28
 DB 5 HGHEQHQHGLGHGKXKLDLLEHQGHV 32

```

RESULT 11
AAY93342
ID AAY93342 standard; protein; 255 AA.
XX
XX
AC AAY93342;
XX
XX
DT 04-SEP-2000 (first entry)
XX
XX
DE Light chain of human high molecular weight kininogen.
XX
XX
KW Human; high molecular weight kininogen; glycoprotein; endothelial cell;
KW Plasma kallikrein; heavy chain; light chain; analogue; angiogenesis;
KW endothelial cell proliferation; endothelial cell migration; vitronectin.
XX
XX
OS Homo sapiens.
XX
XX
PN WO200027415-A2.
XX
XX
PD 18-MAY-2000.
XX
XX
PF 09-NOV-1999; 99WO-US026377.
XX
XX
PR 10-NOV-1998; 98US-0107844P.
XX
XX
PA (UTEM ) UNIV TEMPLE.
PA (DUPO ) DUPONT PHARM CO.
PA (COLM/) COLMAN W R.
PA (MOUS/) MOUSA A S.
XX
XX
PI Colman WR, Mousa AS;
XX
XX
WPI; 2000-376306/32.
XX
XX
DE Method for inhibiting endothelial cell proliferation, using compound that
XX inhibit endothelial cell migration.
XX
XX
PS Disclosure; Page 3; 41pp; English.
XX
XX
CC The present sequence represents the light chain of human high molecular
CC weight kininogen. High molecular weight kininogen is a 120 kDa
CC glycoprotein which binds with high affinity to endothelial cells, where
CC it is cleaved by plasma kallikrein into heavy and light chains. Analogues
CC of high molecular weight kininogen are used in the method of the
CC invention. The specification describes a method of inhibiting endothelial
CC cell proliferation. The method comprises contacting endothelial cells
CC with a compound containing high molecular weight kininogen analogues. The
CC method and the compounds can be used for inhibiting endothelial cell
CC proliferation. The compounds can also be used for inhibiting
CC angiogenesis. The compounds can also be used to inhibit migration of
CC endothelial cells to vitronectin
XX
XX
SQ Sequence 255 AA;

Query Match 98.8%; Score 160; DB 3; Length 255;
Best Local Similarity 96.4%; Pred. No. 9.1e-15;
Matches 27; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 HGHEQHGGLGHGKXKLDLDDLEHGGHV 28
Db 74 HGHEQHGGLGHGKXKLDLDDLEHGGHV 101

RESULT 12
AAB73620
ID AAB73620 standard; protein; 255 AA.
XX
XX
AC AAB73620;
XX
XX
DT 10-AUG-2001 (first entry)
XX
XX
DE Human high molecular weight kininogen (HK) light chain.
XX
XX
KW Human; high molecular weight kininogen; HK; light chain; domain 5;

```

```

KW antibody; monoclonal antibody C11C1; neovascularisation inhibition;
KW endothelial cell proliferation inhibition; antiangiogenic;
KW vascular tube formation inhibition; diabetic retinopathy;
KW rheumatoid arthritis; atherosclerotic plaque rupture; cancer; tumour;
KW cytostatic; antidiabetic; ophthalmological; antirheumatic; antiarthritic;
KW antiatherosclerotic.
XX
XX
OS Homo sapiens.
XX
XX
FH Key Location/Qualifiers
FT Region 1..12 /note= "C-terminal portion of HK domain 4 remaining after
FT /note= "bradykinin liberation"
FT Domain 13..131
FT /label= Domain 5
FT /note= "Corresponds to residues 384-502 of HK"
FT Region 69..131
FT /note= "Referred to in Claim 3; corresponds to HK
FT residues 440-502"
FT Region 69..84
FT /note= "Referred to in Claim 4; corresponds to HK
FT residues 440-457"
FT Domain 70..86
FT /label= His/Gly-rich_subdomain
FT Domain 100..125
FT /label= Endothelial_cell_binding_domain
FT Domain 104..131
FT /label= His/Gly/Lys-rich_subdomain
FT Region 115..131
FT /note= "Referred to in Claim 5; corresponds to HK
FT residues 486-502"
FT Domain 132..255
FT /label= Domain 6
FT /note= "Corresponds to residues 503-626 of HK"
XX
XX
PN WO200134195-A1.
XX
XX
PD 17-MAY-2001.
XX
XX
PR 10-NOV-2000; 2000WO-US030975.
XX
XX
PR 12-NOV-1999; 99US-0165165P.
XX
XX
PA (UTEM ) UNIV TEMPLE.
PA (DUPO ) DUPONT PHARM CO.
XX
XX
PI Colman RW, Mousa SA;
XX
XX
WPI; 2001-328940/34.
XX
XX
DE Inhibiting angiogenesis in a mammal using an antibody against high
XX molecular weight kininogen domain 5.
XX
XX
PS Claim 4; Page 2; 38pp; English.
XX
XX
CC The invention relates to a method of inhibiting angiogenesis in a mammal,
XX comprising administering an antibody against an epitope of high molecular
XX weight kininogen (HK) domain 5. In particular, the antibody used can be
XX monoclonal antibody C11C1 which is produced by hybridoma ATCC HB-8964.
XX The method of the invention is used to inhibit endothelial cell
XX proliferation, vascular tube formation and/or neovascularisation in
XX disease states such as diabetic retinopathy, rheumatoid arthritis and
XX atherosclerotic plaques. The antibody may be administered to prevent
XX plaque rupture, which leads to thrombotic occlusion of coronary or
XX cerebral arteries. The antibody may also be used to inhibit tumour growth
XX via the inhibition of angiogenesis. The present sequence represents human
XX high molecular weight kininogen (HK) light chain, which contains HK
XX domain 5
XX
XX
SQ Sequence 255 AA;

Query Match 98.8%; Score 160; DB 4; Length 255;
Best Local Similarity 96.4%; Pred. No. 9.1e-15;

```

Matches	27;	Conservative	0;	Mismatches	1;	Indels	0;	Gaps	0;
QY	1	HGHEQHQHGLGHGKFKLDDDLHQGGHV	28						
Db	124	HGHEQHQHGLGHGKFKLDDDLHQGGHV	151						
RESULT 14									
ID	ABR41202	standard; protein; 357 AA.							
XX	XX								
AC	ABR41202;								
XX	XX								
DT	02-JUN-2003	(first entry)							
XX	XX								
DE	Human DITHP	extracellular signalling protein.							
XX	XX								
KW	Human; dithp;	diagnostic and therapeutic polynucleotide; diagnosis;							
KW	cancer; cell	proliferative disorder; autoimmune disorder;							
KW	inflammatory	disorder; infection; hormonal disorder; metabolic disorder;							
KW	neurological	disorder; gastrointestinal disorder; transport disorder;							
KW	connective	tissue disorder; drug screening; proteome analysis;							
XX	gene therapy;	antisense therapy; genotyping; transgenic animal; knock in;							
KW	disease model;	toxicological testing; transcript imaging;							
KW	extracellular	signalling.							
XX	XX								
OS	Homo sapiens.								
XX	XX								
PN	W0200297031-A2.								
XX	XX								
PD	05-DEC-2002.								
XX	XX								
PF	27-MAR-2002;	2002WO-US010056.							
XX	XX								
PR	28-MAR-2001;	2001US-0279619P.							
PR	29-MAR-2001;	2001US-0280067P.							
PR	29-MAR-2001;	2001US-0280068P.							
PR	16-MAY-2001;	2001US-0291280P.							
PR	17-MAY-2001;	2001US-0291629P.							
PR	17-MAY-2001;	2001US-0291849P.							
PR	19-JUN-2001;	2001US-0299428P.							
PR	20-JUN-2001;	2001US-0299776P.							
PR	20-JUN-2001;	2001US-0300001P.							
XX	XX								
PA	(INCY-)	INCYTE GENOMICS INC.							
XX	XX								
PI	Daffo A, Jones	AL, Tran AB, Dahl CR, Gietzen D, Chinn J;							
PI	Daifour CB, Hillman	JL, Yu JY, Tuason O, Yap PE, Amshay SR;							
PI	Daughtery SC, Dam	TC, Liu TP, Nguyen DA, Kleefeld Y, Gerstin EH;							
PI	Peralta CH, David	MH, Lewis SA, Chen AJ, Panzer SR, Harris B;							
PI	Flores V, Marwaha	R, Lo A, Ian RY, Urashka ME;							
XX	XX								
DR	WPI; 2003-129518/12.								
DR	N-PSDB; ACC46146.								
XX	XX								
PT	Novel human	diagnostic and therapeutic polypeptide useful for identifying							
PT	test compound	which specifically binds to a polypeptide encoded by human							
PT	diagnostic and	therapeutic polynucleotide, and to induce antibodies.							
XX	XX								
PS	Claim 27; SEQ	ID NO 737; 591pp; English.							
XX	XX								
CC	The invention	relates to novel human diagnostic and therapeutic							
CC	polynucleotides	designated dithp (ACC46080-ACC46749) and to their encoded							
CC	proteins (DITHP;	ABR4136-ABR41812). The invention also relates to							
CC	polynucleotide	sequences at least 90% identical to the dithp cDNA							
CC	sequences of the	invention; recombinant vectors, host cells and							
CC	transgenic	organisms comprising a dithp nucleic acid sequence; the							
CC	recombinant	production of DITHP proteins; antibodies specific for DITHP							
CC	proteins; microarrays	comprising dithp nucleic acid sequences; methods of							
CC	detecting dithp	nucleotide and protein sequences; methods of screening							
CC	for compounds	which specifically bind a DITHP protein; and methods of							
CC	assessing the	toxicity of test compounds using a dithp hybridisation							
CC	probe. Dithp	nucleic acid sequences and DITHP proteins may be used in the							
CC	diagnosis of a	wide variety of conditions including cancer and other cell							
CC	proliferative	disorders; autoimmune or inflammatory disorders; bacterial,							

Matches	27;	Conservative	0;	Mismatches	1;	Indels	0;	Gaps	0;
QY	1	HGHEQHQHGLGHGKFKLDDDLHQGGHV	28						
Db	74	HGHEQHQHGLGHGKFKLDDDLHQGGHV	101						
RESULT 13									
ID	ABG21100	standard; protein; 305 AA.							
XX	XX								
AC	ABG21100;								
XX	XX								
DT	18-FEB-2002	(first entry)							
XX	XX								
DE	Novel human	diagnostic protein #21091.							
XX	XX								
KW	Human; chromosome	mapping; gene mapping; gene therapy; forensic;							
KW	food supplement;	medical imaging; diagnostic; genetic disorder.							
XX	XX								
OS	Homo sapiens.								
XX	XX								
PN	W0200175067-A2.								
XX	XX								
PD	11-OCT-2001.								
XX	XX								
PF	30-MAR-2001;	2001WO-US008631.							
XX	XX								
PR	31-MAR-2000;	2000US-00540217.							
PR	23-AUG-2000;	2000US-00649167.							
XX	XX								
PA	(HYSB-) HYSEQ	INC.							
XX	XX								
PI	Dzmanac RT, Liu	C, Tang YT;							
XX	XX								
DR	WPI; 2001-639362/73.								
XX	XX								
DR	N-PSDB; AAS85287.								
XX	XX								
PT	New isolated	polynucleotide and encoded polypeptides, useful in							
PT	diagnostics, forensic,	gene mapping, identification of mutations							
PT	responsible for	genetic disorders or other traits and to assess							
PT	biodiversity.								
XX	XX								
PS	Claim 20; SEQ	ID NO 51459; 103pp; English.							
XX	XX								
CC	The invention	relates to isolated polynucleotide (I) and polypeptide (II)							
CC	sequences. (I)	is useful as hybridisation probes, polymerase chain							
CC	reaction (PCR)	primers, oligomers, and for chromosome and gene mapping,							
CC	and in recombinant	production of (II). The polynucleotides are also used							
CC	in diagnostics	as expressed sequence tags for identifying expressed							
CC	genes. (I)	is useful in gene therapy techniques to restore normal							
CC	activity of (II)	or to treat disease states involving (II). (II) is							
CC	useful for	generating antibodies against it, detecting or quantitating a							
CC	polypeptide in	tissue, as molecular weight markers and as a food							
CC	supplement. (II)	and its binding partners are useful in medical imaging							
CC	of sites	expressing (II). (I) and (II) are useful for treating disorders							
CC	involving	aberrant protein expression or biological activity. The							
CC	polypeptide and	polynucleotide sequences have applications in							
CC	diagnostics, forensic,	gene mapping, identification of mutations							
CC	responsible for	genetic disorders or other traits to assess biodiversity							
CC	and to produce	other types of data and products dependent on DNA and							
CC	amino acid	sequences. ABG00010-ABG30377 represent novel human diagnostic							
CC	amino acid	sequences of the invention. Note: The sequence data for this							
CC	patent did not	appear in the printed specification, but was obtained in							
CC	electronic	format directly from WIPO at							
CC	ftp.wipo.int/pub/published	pct_sequences							
XX	XX								
SQ	Sequence	305 AA;							
Query Match	98.8%;	Score 160;	DB 4;	Length 305;					
Best Local Similarity	96.4%;	Pred. No. 1.1e-14;							
Matches	27;	Conservative	0;	Mismatches	1;	Indels	0;	Gaps	0;

CC viral, fungal or parasitic infections; hormonal disorders; metabolic
 CC disorders; neurological disorders; gastrointestinal disorders; transport
 CC disorders; and connective tissue disorders. They may also be used to
 CC screen for modulators of protein activity or gene expression. DITHP
 CC proteins can additionally be used in analysis of the proteome of a tissue
 CC or cell type and to induce antibodies. The dithp nucleic acids are
 CC additionally useful in somatic or germline gene therapy of the disorders
 CC mentioned above, as a source of antisense sequences, as a source of
 CC probes and primers, in genotyping and identification of individuals, in
 CC the generation of transgenic animal models of human disease or knock in
 CC humanised animals, in toxicological testing, and in transcript imaging.
 CC The present sequence represents a DITHP protein which has extracellular
 CC part of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 357 AA;
 Query Match 98.8%; Score 160; DB 6; Length 357;
 Best Local Similarity 96.4%; Pred. No. 1.3e-14;
 Matches 27; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 HGHEQQRGLGHHGKXKLDLLEHQQGHV 28
 DB 176 HGHEQQRGLGHHGKXKLDLLEHQQGHV 203
 RESULT 15
 ABU99146
 ID ABU99146 standard; protein; 415 AA.
 XX
 AC ABU99146;
 XX
 DT 01-AUG-2003 (first entry)
 XX
 DE Novel human GPCR related protein NOV12d.
 XX
 KW Human; G-protein coupled receptor related protein; GPCR related protein;
 KW NOV; cytosolic; cardiant; antiarteriosclerotic; antidiabetic;
 KW immunomodulator; anti-HIV; anorectic; antidiabetic; haemostatic;
 KW antiparkinsonian; neuroprotective; nootropic; gene therapy; vaccine;
 KW NOVX-associated disorder; cardiomyopathy; atherosclerosis; cancer;
 KW diabetes; immune disorder; AIDS; obesity; asthma;
 KW haematopoietic disorder; Parkinson's disease; Alzheimer's disease;
 KW infection; multiple sclerosis; cancer-associated cachexia;
 KW wasting disorder; chronic disease; neurogenesis; cell differentiation;
 KW cell proliferation; haematopoiesis; wound healing; angiogenesis;
 KW chromosome mapping; tissue typing; preventive medicine; pharmacogenomic.
 XX
 OS Homo sapiens.
 XX
 PN WO200299116-A2.
 XX
 PD 12-DEC-2002.
 XX
 PF 04-JUN-2002; 2002WO-US017428.
 XX
 PR 04-JUN-2001; 2001US-0295607P.
 PR 04-JUN-2001; 2001US-0295661P.
 PR 06-JUN-2001; 2001US-0296404P.
 PR 06-JUN-2001; 2001US-0296418P.
 PR 14-JUN-2001; 2001US-0298285P.
 PR 15-JUN-2001; 2001US-0298566P.
 PR 21-JUN-2001; 2001US-0299949P.
 PR 26-JUN-2001; 2001US-0300883P.
 PR 28-JUN-2001; 2001US-0301550P.
 PR 13-AUG-2001; 2001US-0311972P.
 PR 27-AUG-2001; 2001US-0315071P.
 PR 29-AUG-2001; 2001US-0315669P.
 PR 14-SEP-2001; 2001US-0322293P.
 PR 17-SEP-2001; 2001US-0322706P.
 PR 14-DEC-2001; 2001US-0341186P.
 PR 28-FEB-2002; 2002US-0361189P.

PR 12-MAR-2002; 2002US-0363673P.
 PR 12-MAR-2002; 2002US-0363676P.
 XX 03-JUN-2002; 2002US-00363676.
 XX (CURA-) CUPAGEN CORP.
 PA Anderson DW, Baumgartner JC, Boldog FL, Casman SJ, Edinger SR;
 PI Gangolli EA, Gerlach VL, Gorman L, Guo X, Hjalt T, Kekuda R, Li L;
 PI Macdougall JR, Malyanankar DV, Millet I, Padigara M, Patturajan M;
 PI Pena CEA, Raetelli L, Shinkets RA, Stone DJ, Spytek KA, Vernet CAM;
 PI Voss EZ, Zerhusen BD;
 XX WPI; 2003-140627/13.
 DR N-PSDB; ACD03650.
 XX New NOVX polypeptides and nucleic acids, useful for preventing or
 PT treating NOVX-associated disorders, e.g. cancer, cardiomyopathy,
 PT atherosclerosis, or diabetes, and in chromosome mapping, tissue typing or
 PT pharmacogenomics.
 XX Claim 1; Page 145; 332pp; English.
 XX The invention describes an isolated polypeptide (I) comprising any of 27
 CC 118-961 residue amino acid sequences, given in the specification, a
 CC mature form of them, a sequence that is at least 95 % identical to them,
 CC or a sequence having one or more conservative substitutions in them. The
 CC polypeptide is useful in manufacturing a medicament for treating a
 CC syndrome associated with a human disease selected from a pathology
 CC associated with the polypeptide. The NOVX polypeptides, polynucleotides
 CC and antibodies are useful in treating or preventing NOVX-associated
 CC disorders, e.g. cardiomyopathy, atherosclerosis, cancer, diabetes, immune
 CC disorders, AIDS, obesity, asthma, haematopoietic disorders, Parkinson's
 CC disease, Alzheimer's disease, infections, multiple sclerosis, cancer-
 CC associated cachexia, and other wasting disorders associated with chronic
 CC diseases. The nucleic acids and polypeptides may also be used as targets
 CC for the identification of small molecules that modulate or inhibit e.g.
 CC neurogenesis, cell differentiation, cell proliferation, haematopoiesis,
 CC wound healing and angiogenesis, in gene therapy, in generation of
 CC antibodies that bind immunospecifically to NOVX substances for use in
 CC therapeutic or diagnostic methods. The nucleic acids are further used as
 CC hybridisation probes, in chromosome mapping, tissue typing, preventive
 CC medicine, and pharmacogenomics. The polypeptides are also useful as
 CC vaccines. This is the amino acid sequence of a novel human G-protein
 CC coupled receptor related protein NOV
 XX Sequence 415 AA;
 SQ
 Query Match 98.8%; Score 160; DB 6; Length 415;
 Best Local Similarity 96.4%; Pred. No. 1.5e-14;
 Matches 27; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 HGHEQQRGLGHHGKXKLDLLEHQQGHV 28
 DB 234 HGHEQQRGLGHHGKXKLDLLEHQQGHV 261
 Search completed: March 4, 2004, 13:07:26
 Job time : 77.8235 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2004, 13:04:08 ; Search time 21.4118 Seconds
(without alignments)
67.511 Million cell updates/sec

Title: SEQ-B

Perfect score: 162

Sequence: 1 HGHEQHGGLGHGKXKLDLLEHOGGHV 28

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:
1: /cgn2_6/ptodata/2/iaa/5A COMB.pcp.*
2: /cgn2_6/ptodata/2/iaa/5B COMB.pcp.*
3: /cgn2_6/ptodata/2/iaa/6A COMB.pcp.*
4: /cgn2_6/ptodata/2/iaa/6B COMB.pcp.*
5: /cgn2_6/ptodata/2/iaa/PTCUS COMB.pcp.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	160	98.8	55	3	US-09-612-126-5
2	160	98.8	62	3	US-09-612-126-7
3	160	98.8	83	3	US-09-612-126-6
4	160	98.8	94	3	US-09-612-126-10
5	160	98.8	186	3	US-09-612-126-8
6	160	98.8	255	3	US-09-612-126-1
7	138	85.2	179	3	US-09-612-126-11
8	91	56.2	47	3	US-09-612-126-4
9	69	42.6	15	3	US-09-612-126-2
10	69	42.6	16	3	US-09-612-126-9
11	68.5	42.3	474	3	US-09-461-474-10
12	66	40.7	344	4	US-09-134-001C-3524
13	64.5	39.8	448	3	US-09-461-474-8
14	61	37.7	21	3	US-09-612-126-3
15	60.5	37.3	398	3	US-09-461-474-17
16	59	36.4	715	4	US-09-252-991A-19235
17	57	35.2	313	3	US-09-456-287-3
18	57	35.2	333	3	US-08-686-528A-2
19	57	35.2	337	3	US-09-456-287-2
20	57	35.2	337	3	US-09-456-287-2
21	57	35.2	535	4	US-09-252-991A-28410
22	55	34.0	680	4	US-09-252-991A-29223
23	55	34.0	788	4	US-09-252-991A-17380
24	54	33.3	339	4	US-09-328-352-6551
25	54	33.3	339	4	US-09-328-352-6551
26	53.5	33.0	270	4	US-09-252-991A-29237
27	52.5	32.4	370	4	US-09-252-991A-20626
					Sequence 23852, A

28	52.5	32.4	572	4	US-09-252-991A-17898	Sequence 17898, A
29	52	32.1	1174	2	US-08-446-345-36	Sequence 36, Appl
30	52	32.1	1958	1	US-07-945-283-2	Sequence 2, Appl
31	51.5	31.8	226	4	US-09-489-039A-10824	Sequence 10824, A
32	51.5	31.8	249	4	US-09-543-681A-6793	Sequence 6793, Ap
33	51.5	31.8	344	4	US-09-252-991A-23772	Sequence 23772, A
34	51.5	31.8	378	4	US-09-252-991A-22754	Sequence 22754, A
35	51.5	31.8	856	4	US-09-252-991A-17850	Sequence 17850, A
36	51	31.5	762	4	US-09-252-991A-22060	Sequence 22060, A
37	50.5	31.2	160	4	US-09-252-991A-19037	Sequence 19037, A
38	50.5	31.2	364	4	US-09-252-991A-22803	Sequence 22803, A
39	50.5	31.2	496	4	US-09-252-991A-21627	Sequence 21627, A
40	50.5	31.2	619	4	US-09-543-681A-6151	Sequence 6151, Ap
41	50	30.9	400	4	US-09-107-532A-4501	Sequence 4501, Ap
42	50	30.9	414	4	US-09-107-532A-4501	Sequence 23442, A
43	50	30.9	469	4	US-09-252-991A-28058	Sequence 28058, A
44	50	30.9	1049	4	US-09-252-991A-28058	Sequence 14, Appl
45	49.5	30.6	423	4	US-09-668-262A-14	

ALIGNMENTS

RESULT 1
US-09-612-126-5
; Sequence 5, Application US/09612126
; Patent No. 6284726
; GENERAL INFORMATION:
; APPLICANT: TEMPLE UNIVERSITY - OF THE COMMONWEALTH SYSTEM OF
; TITLE OF INVENTION: INHIBITION OF ANGIOGENESIS BY PEPTIDE ANALOGS OF HIGH
; FILE REFERENCE: 6056-258 CTI
; CURRENT APPLICATION NUMBER: US/09/612,126
; CURRENT FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/107,844
; PRIOR FILING DATE: 1998-11-10
; PRIOR APPLICATION NUMBER: PCT/US99/26377
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 55
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Human high
; OTHER INFORMATION: molecular weight kininogen light chain amino acids
; OTHER INFORMATION: Lys(420) through Asp(474)
; US-09-612-126-5

Query Match 98.8%; Score 160; DB 3; Length 55;
Best Local Similarity 96.4%; Pred. No. 1.3e-15;
Matches 27; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 HGHEQHGGLGHGKXKLDLLEHOGGHV 28
Db 26 HGHEQHGGLGHGKXKLDLLEHOGGHV 53

RESULT 2
US-09-612-126-7
; Sequence 7, Application US/09612126
; Patent No. 6284726
; GENERAL INFORMATION:
; APPLICANT: TEMPLE UNIVERSITY - OF THE COMMONWEALTH SYSTEM OF
; TITLE OF INVENTION: INHIBITION OF ANGIOGENESIS BY PEPTIDE ANALOGS OF HIGH
; FILE REFERENCE: 6056-258 CTI
; CURRENT APPLICATION NUMBER: US/09/612,126
; CURRENT FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/107,844
; PRIOR FILING DATE: 1998-11-10
; PRIOR APPLICATION NUMBER: PCT/US99/26377

THE UNIVERSITY OF CHICAGO PRESS

```

; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 255
; TYPE: PRT
; ORGANISM: Human
US-09-612-126-1

Query Match      98.8%; Score 160; DB 3; Length 255;
Best Local Similarity 96.4%; Pred. No. 7e-15;
Matches 27; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 HGHEQOGLGHGKXKLDLLEHQGHV 28
        ||||| ||||| ||||| ||||| |||||
Db      74 HGHEQOGLGHGKXKLDLLEHQGHV 101

RESULT 7
US-09-612-126-11
; Sequence 11, Application US/09612126
; Patent No. 6284726
; GENERAL INFORMATION:
; APPLICANT: TEMPLE UNIVERSITY - OF THE COMMONWEALTH SYSTEM OF
; TITLE OF INVENTION: INHIBITION OF ANGIOGENESIS BY PEPTIDE ANALOGS OF HIGH
; TITLE OF INVENTION: MOLECULAR WEIGHT KININOGEN DOMAN 5
; FILE REFERENCE: 6056-258 CT1
; CURRENT APPLICATION NUMBER: US/09/612,126
; CURRENT FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/107,844
; PRIOR FILING DATE: 1998-11-10
; PRIOR APPLICATION NUMBER: PCT/US99/26377
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 179
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Human high
; OTHER INFORMATION: molecular weight kininogen light chain amino acids
; OTHER INFORMATION: Glu(448) through Ser(626)
US-09-612-126-11

Query Match      85.2%; Score 138; DB 3; Length 179;
Best Local Similarity 96.0%; Pred. No. 5.5e-12;
Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      4 EQQGLGHGKXKLDLLEHQGHV 28
        ||||| ||||| ||||| ||||| |||||
Db      1 EQQGLGHGKXKLDLLEHQGHV 25

RESULT 8
US-09-612-126-4
; Sequence 4, Application US/09612126
; Patent No. 6284726
; GENERAL INFORMATION:
; APPLICANT: TEMPLE UNIVERSITY - OF THE COMMONWEALTH SYSTEM OF
; TITLE OF INVENTION: INHIBITION OF ANGIOGENESIS BY PEPTIDE ANALOGS OF HIGH
; TITLE OF INVENTION: MOLECULAR WEIGHT KININOGEN DOMAN 5
; FILE REFERENCE: 6056-258 CT1
; CURRENT APPLICATION NUMBER: US/09/612,126
; CURRENT FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/107,844
; PRIOR FILING DATE: 1998-11-10
; PRIOR APPLICATION NUMBER: PCT/US99/26377
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 47
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:

```

```

; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Human high
; OTHER INFORMATION: molecular weight kininogen light chain amino acids
; OTHER INFORMATION: Gly(456) through Lys(502)
US-09-612-126-4

Query Match      56.2%; Score 91; DB 3; Length 47;
Best Local Similarity 94.1%; Pred. No. 4.4e-06;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      12 GHKXKLDLLEHQGHV 28
        ||||| ||||| ||||| ||||| |||||
Db      1 GHKFKLDLLEHQGHV 17

RESULT 9
US-09-612-126-2
; Sequence 2, Application US/09612126
; Patent No. 6284726
; GENERAL INFORMATION:
; APPLICANT: TEMPLE UNIVERSITY - OF THE COMMONWEALTH SYSTEM OF
; TITLE OF INVENTION: INHIBITION OF ANGIOGENESIS BY PEPTIDE ANALOGS OF HIGH
; TITLE OF INVENTION: MOLECULAR WEIGHT KININOGEN DOMAN 5
; FILE REFERENCE: 6056-258 CT1
; CURRENT APPLICATION NUMBER: US/09/612,126
; CURRENT FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/107,844
; PRIOR FILING DATE: 1998-11-10
; PRIOR APPLICATION NUMBER: PCT/US99/26377
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Human high
; OTHER INFORMATION: molecular weight kininogen light chain amino acids
; OTHER INFORMATION: His (441) through His (455)
US-09-612-126-2

Query Match      42.6%; Score 69; DB 3; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.0015;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 HGHEQOGLGH 11
        ||||| ||||| |||||
Db      5 HGHEQOGLGH 15

RESULT 10
US-09-612-126-9
; Sequence 9, Application US/09612126
; Patent No. 6284726
; GENERAL INFORMATION:
; APPLICANT: TEMPLE UNIVERSITY - OF THE COMMONWEALTH SYSTEM OF
; TITLE OF INVENTION: INHIBITION OF ANGIOGENESIS BY PEPTIDE ANALOGS OF HIGH
; TITLE OF INVENTION: MOLECULAR WEIGHT KININOGEN DOMAN 5
; FILE REFERENCE: 6056-258 CT1
; CURRENT APPLICATION NUMBER: US/09/612,126
; CURRENT FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/107,844
; PRIOR FILING DATE: 1998-11-10
; PRIOR APPLICATION NUMBER: PCT/US99/26377
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:

```


seq-b.ra1

Fri Mar 5 08:53:30 2004

Query Match 42.6%; Score 69; DB 3; Length 16;
Best Local Similarity 100.0%; Pred. No. 0.0016;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OTHER INFORMATION: Description of Artificial Sequence: Human high
molecular weight kininogen light chain amino acids
OTHER INFORMATION: Gly(440) through His(455)
US-09-612-126-9

Query Match 42.6%; Score 69; DB 3; Length 16;
Best Local Similarity 100.0%; Pred. No. 0.0016;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HGEHQGHG 11
DB 6 HGEHQGHG 16

RESULT 11
US-09-461-474-10
Sequence 10, Application US/09461474
Patent No. 6278042

GENERAL INFORMATION:
APPLICANT: Allen, Steve
APPLICANT: Rafalski, Antoni
APPLICANT: Sakai, Hajime
TITLE OF INVENTION: Plant Metal Transporters
FILE REFERENCE: BE1303 US NA
CURRENT APPLICATION NUMBER: US/09/461,474
CURRENT FILING DATE: 1999-12-14
EARLIER APPLICATION NUMBER: 60/112,562
EARLIER FILING DATE: 1998-12-16
NUMBER OF SEQ ID NOS: 17
SOFTWARE: Microsoft Office 97
SEQ ID NO 10
TYPE: PRT
LENGTH: 474
ORGANISM: Oryza sativa
US-09-461-474-10

Query Match 42.3%; Score 68.5; DB 3; Length 474;
Best Local Similarity 41.0%; Pred. No. 0.075; Indels 13; Gaps 2;
Matches 16; Conservative 1; Mismatches 9; Indels 13; Gaps 2;

QY 1 HGEHQGHG 28
DB 240 HGEHQGHG 276

RESULT 12
US-09-134-001C-3524
Sequence 3524, Application US/09134001C
Patent No. 6380370

GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
CURRENT FILING DATE: 1998-08-13

PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 3524
LENGTH: 344
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3524

Query Match 40.7%; Score 66; DB 4; Length 344;
Best Local Similarity 44.0%; Pred. No. 0.12; Indels 10; Gaps 0;
Matches 11; Conservative 4; Mismatches 10; Indels 10; Gaps 0;

QY 3 HGEHQGHG 27
DB 11 HGEHQGHG 27

Db 144 HHEHGEHGHESHEDHEHEHHH 168

RESULT 13
US-09-461-474-8
Sequence 8, Application US/09461474
Patent No. 6278042

GENERAL INFORMATION:
APPLICANT: Allen, Steve
APPLICANT: Rafalski, Antoni
APPLICANT: Sakai, Hajime
TITLE OF INVENTION: Plant Metal Transporters
FILE REFERENCE: BE1303 US NA
CURRENT APPLICATION NUMBER: US/09/461,474
CURRENT FILING DATE: 1999-12-14
EARLIER APPLICATION NUMBER: 60/112,562
EARLIER FILING DATE: 1998-12-16
NUMBER OF SEQ ID NOS: 17
SOFTWARE: Microsoft Office 97
SEQ ID NO 8
LENGTH: 448
TYPE: PRT
ORGANISM: Zea mays
US-09-461-474-8

Query Match 39.8%; Score 64.5; DB 3; Length 448;
Best Local Similarity 36.8%; Pred. No. 0.257; Indels 11; Gaps 2;
Matches 14; Conservative 3; Mismatches 10; Indels 11; Gaps 2;

QY 1 HGEHQGHG 27
DB 226 HGEHQGHG 263

RESULT 14
US-09-612-126-3
Sequence 3, Application US/09612126
Patent No. 6284726

GENERAL INFORMATION:
APPLICANT: TEMPLE UNIVERSITY - OF THE COMMONWEALTH SYSTEM OF
TITLE OF INVENTION: INHIBITION OF ANGIOGENESIS BY PEPTIDE ANALOGS OF HIGH
MOLECULAR WEIGHT KININOGEN DOMAN 5
FILE REFERENCE: 6056-258 CT1
CURRENT APPLICATION NUMBER: US/09/612,126
CURRENT FILING DATE: 2000-07-07

PRIOR APPLICATION NUMBER: 60/107,844
PRIOR FILING DATE: 1998-11-10
PRIOR APPLICATION NUMBER: PCT/US99/26377
PRIOR FILING DATE: 1999-11-09
NUMBER OF SEQ ID NOS: 12
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 3
LENGTH: 21
TYPE: PRT
ORGANISM: Artificial Sequence

FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Human high
molecular weight kininogen light chain amino acids
OTHER INFORMATION: Lys(420) through Gly(440)
US-09-612-126-3

Query Match 37.7%; Score 61; DB 3; Length 21;
Best Local Similarity 52.4%; Pred. No. 0.028; Indels 8; Gaps 0;
Matches 11; Conservative 2; Mismatches 8; Indels 8; Gaps 0;

QY 6 HGEHQGHG 26
DB 1 HGEHQGHG 21

RESULT 15
US-09-461-474-17
Sequence 17, Application US/09461474

seq-b.rai

Fri Mar 5 08:53:30 2004

```

; Patent No. 6278042
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Rafalski, Antoni
; APPLICANT: Sakai, Hajime
; TITLE OF INVENTION: Plant Metal Transporters
; FILE REFERENCE: BB1303 US NA
; CURRENT APPLICATION NUMBER: US/09/461,474
; CURRENT FILING DATE: 1999-12-14
; EARLIER APPLICATION NUMBER: 60/112,562
; EARLIER FILING DATE: 1998-12-16
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 17
; LENGTH: 398
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; US-09-461-474-17

Query Match      37.3%; Score 60.5; DB 3; Length 398;
Best Local Similarity 46.2%; Pred. No. 0.8;
Matches 12; Conservative 0; Mismatches 9; Indels 5; Gaps 1;

Qy      1 HGHGQGHGHLGHGKXKLDLLEHGG 26
      |||||
Db     184 HGHSHGHGHHGHL-----DHNHSHG 204

Search completed: March 4, 2004, 13:12:26
Job time : 21.4118 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2004, 13:10:38 ; Search time 42.4118 Seconds
(without alignments)
139.402 Million cell updates/sec

Title: SEQ-B

Perfect score: 162

Sequence: 1 HGHEQHQHGHGHHKXLDLDLHGHGHV 28

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 809742 seqs, 211153259 residues

Total number of hits satisfying chosen parameters: 809742

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubaa/US09_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/2/pubaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	160	98.8	415	15	US-10-162-335-76
2	160	98.8	615	15	US-10-162-335-72
3	160	98.8	644	15	US-10-162-335-74
4	160	98.8	644	15	US-10-162-335-84
5	72	44.4	574	14	US-10-156-761-14106
6	68.5	42.3	378	14	US-10-176-306-59
7	64	39.5	52	13	US-10-090-035-24
8	63.5	39.2	376	14	US-10-176-306-29
9	63.5	39.2	376	15	US-10-264-237-1632
10	63.5	39.2	476	15	US-10-360-849A-33
11	62	38.3	385	14	US-10-270-333-42
12	61	37.7	92	13	US-10-090-035-20
13	59	36.4	91	13	US-10-090-035-22
14	58.5	36.1	469	15	US-10-360-849A-36
15	57	35.2	213	14	US-10-156-761-12186

16	57	35.2	774	14	US-10-156-761-11628
17	56	34.6	377	16	US-10-380-727-6
18	55	34.0	245	15	US-10-369-493-6977
19	55	34.0	444	15	US-10-360-849A-30
20	54	33.3	503	9	US-09-738-897-5
21	54	33.3	507	9	US-09-738-897-4
22	54	33.3	507	10	US-09-849-138-41
23	54	33.3	507	14	US-10-176-306-49
24	53	32.7	831	9	US-09-789-561-86
25	53	32.7	831	11	US-09-833-245-2153
26	51	31.5	92	14	US-10-029-386-29812
27	50.5	31.2	130	16	US-10-389-566-2423
28	50.5	31.2	133	16	US-10-389-566-2424
29	50.5	31.2	261	16	US-10-375-693-28
30	50.5	31.2	265	16	US-10-389-566-2422
31	50.5	31.2	824	14	US-10-156-761-12970
32	50.5	31.2	19695	15	US-10-084-846A-3
33	49.5	30.6	423	14	US-10-427-442-14
34	49	30.2	69	14	US-10-156-761-8065
35	49	30.2	114	15	US-10-312-273-359
36	49	30.2	136	15	US-10-289-762-499
37	49	30.2	476	15	US-10-264-049-2300
38	49	30.2	856	10	US-09-948-029-66
39	48	29.6	60	10	US-09-820-843A-13
40	48	29.6	105	9	US-09-925-295-1231
41	48	29.6	105	10	US-09-925-295-1231
42	48	29.6	339	10	US-09-934-455-492
43	48	29.6	417	15	US-10-116-275-224
44	48	29.6	968	15	US-10-291-172-739
45	47.5	29.3	77	9	US-09-925-302-875

ALIGNMENTS

RESULT 1

- US-10-162-335-76
- ; Sequence 76, Application US/10162335
- ; Publication No. US20040009480A1
- ; GENERAL INFORMATION:
- ; APPLICANT: Anderson, David W.
- ; APPLICANT: Baumgartner, Jason C.
- ; APPLICANT: Boldog, Ferenc L.
- ; APPLICANT: Casman, Stacie J.
- ; APPLICANT: Edinger, Shlomit R.
- ; APPLICANT: Gangolli, Esha A.
- ; APPLICANT: Gerlach, Valerie
- ; APPLICANT: Gorman, Linda
- ; APPLICANT: Guo, Xiaojia (Sasha)
- ; APPLICANT: Hjal, Tord
- ; APPLICANT: Kekuda, Ramesh
- ; APPLICANT: Li, Li
- ; APPLICANT: MacDougall, John R.
- ; APPLICANT: Malyankar, Uriel M.
- ; APPLICANT: Millet, Isabelle
- ; APPLICANT: Padigaru, Muralidhara
- ; APPLICANT: Patturajan, Meera
- ; APPLICANT: Pena, Carol E. A.
- ; APPLICANT: Rastelli, Luca
- ; APPLICANT: Shinkets, Richard A.
- ; APPLICANT: Stone, David J.
- ; APPLICANT: Spytek, Kimberly A.
- ; APPLICANT: Vernet, Corine A. M.
- ; APPLICANT: Vosse, Edward Z.
- ; APPLICANT: Zerhusen, Bryan D.
- ; TITLE OF INVENTION: Therapeutic Polypeptides, Nucleic Acids Encoding Same, and Methods
- ; FILE REFERENCE: 21402-377 B
- ; CURRENT APPLICATION NUMBER: US/10162,335
- ; CURRENT FILING DATE: 2002-10-01
- ; PRIOR APPLICATION NUMBER: 60/295,607
- ; PRIOR FILING DATE: 2001-06-04
- ; PRIOR APPLICATION NUMBER: 60/295,661
- ; PRIOR FILING DATE: 2001-06-04

APPLICANT: Zerhusen, Bryan D.

; PRIOR APPLICATION NUMBER: 60/297,414
; PRIOR FILING DATE: 2001-06-11
; PRIOR APPLICATION NUMBER: 60/297,567
; PRIOR FILING DATE: 2001-06-12
; PRIOR APPLICATION NUMBER: 60/298,285
; PRIOR FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: 60/298,556
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/299,949
; PRIOR FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: 60/300,883
; PRIOR FILING DATE: 2001-06-26
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 201
; SEQ ID NO 74
; LENGTH: 644
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-162-335-74

Query Match 98.8%; Score 160; DB 15; Length 644;
Best Local Similarity 96.4%; Pred. No. 4.7e-13;
Matches 27; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 HGHEQOHGLGHGKXKLDLDDLEHOGGHV 28
DB 463 HGHEQOHGLGHGKXKLDLDDLEHOGGHV 490

RESULT 4

US-10-162-335-84
; Sequence 84, Application US/10162335
; Publication No. US20040009480A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, David W.
; APPLICANT: Baumgartner, Jason C.
; APPLICANT: Boldog, Ferenc L.
; APPLICANT: Casman, Stacie J.
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Gangolli, Bsha A.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Gorman, Linda
; APPLICANT: Guo, Xiaojia (Sasha)
; APPLICANT: Hjalt, Tord
; APPLICANT: Kexuda, Ramesh
; APPLICANT: Li, Li
; APPLICANT: MacDougall, John R.
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Millet, Isabelle
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Patturajan, Meera
; APPLICANT: Pena, Carol E. A.
; APPLICANT: Rastelli, Luca
; APPLICANT: Shinkets, Richard A.
; APPLICANT: Stone, David J.
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Vernet, Corine A. M.
; APPLICANT: Voss, Edward Z.
; APPLICANT: Zethuesen, Bryan D.
; TITLE OF INVENTION: Therapeutic Polypeptides, Nucleic Acids Encoding Same, and Method
; FILE REFERENCE: 21402-377 B
; CURRENT APPLICATION NUMBER: US/10/162,335
; CURRENT FILING DATE: 2002-10-01
; PRIOR APPLICATION NUMBER: 60/295,607
; PRIOR FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: 60/295,661
; PRIOR FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: 60/296,404
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: 60/296,418
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: 60/297,414
; PRIOR FILING DATE: 2001-06-11

; PRIOR APPLICATION NUMBER: 60/297,567
; PRIOR FILING DATE: 2001-06-12
; PRIOR APPLICATION NUMBER: 60/298,285
; PRIOR FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: 60/298,556
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/299,949
; PRIOR FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: 60/300,883
; PRIOR FILING DATE: 2001-06-26
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 201
; SEQ ID NO 84
; LENGTH: 644
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-162-335-84

Query Match 98.8%; Score 160; DB 15; Length 644;
Best Local Similarity 96.4%; Pred. No. 4.7e-13;
Matches 27; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 HGHEQOHGLGHGKXKLDLDDLEHOGGHV 28
DB 463 HGHEQOHGLGHGKXKLDLDDLEHOGGHV 490

RESULT 5

US-10-156-761-14106
; Sequence 14106, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 14106
; LENGTH: 574
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-14106

Query Match 44.4%; Score 72; DB 14; Length 574;
Best Local Similarity 52.2%; Pred. No. 0.18;
Matches 12; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

QY 1 HGHEQOHGLGHGKXKLDLDDLEH 23
DB 418 HXENHGHGHGKXKLDLDDLEH 440

RESULT 6

US-10-176-306-59
; Sequence 59, Application US/10176306
; Publication No. US20030130485A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel A.
; APPLICANT: Curtis, Rory A. J.
; APPLICANT: Glucksmann, Maria Alexandra
; APPLICANT: Bandaru, Rajasekhar
; APPLICANT: Kapeller-Libermann, Rosana

seq-b.rapb

Fri Mar 5 08:53:30 2004

; TITLE OF INVENTION: NOVEL HUMAN GENES AND METHODS OF USE THEREOF

; FILE REFERENCE: 10448-195001

; CURRENT APPLICATION NUMBER: US/10/176,306

; CURRENT FILING DATE: 2002-06-20

; PRIOR APPLICATION NUMBER: 10/001,137

; PRIOR FILING DATE: 2001-11-14

; PRIOR APPLICATION NUMBER: PCT/US01/45291

; PRIOR FILING DATE: 2001-11-14

; PRIOR APPLICATION NUMBER: 60/248,362

; PRIOR FILING DATE: 2000-11-14

; PRIOR APPLICATION NUMBER: 60/248,331

; PRIOR FILING DATE: 2000-11-14

; PRIOR APPLICATION NUMBER: 60/248,365

; PRIOR FILING DATE: 2000-11-14

; PRIOR APPLICATION NUMBER: 60/250,077

; PRIOR FILING DATE: 2000-11-14

; PRIOR APPLICATION NUMBER: 60/250,327

; PRIOR FILING DATE: 2000-11-30

; PRIOR APPLICATION NUMBER: 60/250,176

; PRIOR FILING DATE: 2000-11-30

; PRIOR APPLICATION NUMBER: 10/023,617

; PRIOR FILING DATE: 2001-12-18

; PRIOR APPLICATION NUMBER: PCT/US01/49416

; PRIOR FILING DATE: 2001-12-18

; PRIOR APPLICATION NUMBER: 60/256,249

; PRIOR FILING DATE: 2000-12-18

; PRIOR APPLICATION NUMBER: 60/256,405

; PRIOR FILING DATE: 2000-12-18

; PRIOR APPLICATION NUMBER: 10/083,248

; PRIOR FILING DATE: 2001-10-22

; PRIOR APPLICATION NUMBER: PCT/US01/46717

; PRIOR FILING DATE: 2001-10-22

; PRIOR APPLICATION NUMBER: 60/242,324

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/242,518

; PRIOR FILING DATE: 2000-10-23

; PRIOR APPLICATION NUMBER: 60/241,969

; PRIOR FILING DATE: 2000-10-20

; NUMBER OF SEQ ID NOS: 86

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 59

; TYPE: PRT

; ORGANISM: Mus musculus

US-10-176-306-59

Query Match 42.3%; Score 68.5; DB 14; Length 378;

Best Local Similarity 48.1%; Pred. No. 0.33;

Matches 13; Conservative 2; Mismatches 11; Indels 1; Gaps 1;

Qy 2 GHEQOHGLGHGKXKL-DDDLEHOGG 27

Db 163 GHGHHGSHGSHSLFNGALDHSHG 189

RESULT 7

US-10-090-035-24

; Sequence 24; Application US/10090035

; Publication No. US20020170089A1

; GENERAL INFORMATION:

; APPLICANT: Simmons, Carl R.

; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible

; FILE REFERENCE: 35718/242990

; CURRENT APPLICATION NUMBER: US/10/090,035

; CURRENT FILING DATE: 2002-02-28

; PRIOR APPLICATION NUMBER: 60/272,227

; PRIOR FILING DATE: 02/28/2001

; NUMBER OF SEQ ID NOS: 25

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 24

; LENGTH: 92

; TYPE: PRT

; ORGANISM: Triticum aestivum

US-10-090-035-24

Query Match 39.5%; Score 64; DB 13; Length 92;

Best Local Similarity 40.0%; Pred. No. 0.3;

Matches 12; Conservative 5; Mismatches 9; Indels 4; Gaps 1;

Qy 1 HGHEQHGLGHGKXKLDDLEHOGG 26

Db 48 HGHEQHGLGHGSHSLFNGALDHSHG 77

RESULT 8

US-10-176-306-29

; Sequence 29; Application US/10176306

; Publication No. US20030130485A1

; GENERAL INFORMATION:

; APPLICANT: Meyers, Rachel A.

; APPLICANT: Glucksmann, Maria Alexandra

; APPLICANT: Bandaru, Rajasekhar

; APPLICANT: Kapeller-Libermann, Rosana

; TITLE OF INVENTION: NOVEL HUMAN GENES AND METHODS OF USE THEREOF

; FILE REFERENCE: 10448-195001

; CURRENT APPLICATION NUMBER: US/10/176,306

; CURRENT FILING DATE: 2002-06-20

; PRIOR APPLICATION NUMBER: 10/001,137

; PRIOR FILING DATE: 2001-11-14

; PRIOR APPLICATION NUMBER: PCT/US01/45291

; PRIOR FILING DATE: 2001-11-14

; PRIOR APPLICATION NUMBER: 60/248,362

; PRIOR FILING DATE: 2000-11-14

; PRIOR APPLICATION NUMBER: 60/248,331

; PRIOR FILING DATE: 2000-11-14

; PRIOR APPLICATION NUMBER: 60/248,365

; PRIOR FILING DATE: 2000-11-14

; PRIOR APPLICATION NUMBER: 60/250,077

; PRIOR FILING DATE: 2000-11-30

; PRIOR APPLICATION NUMBER: 60/250,327

; PRIOR FILING DATE: 2000-11-30

; PRIOR APPLICATION NUMBER: 60/250,176

; PRIOR FILING DATE: 2000-11-30

; PRIOR APPLICATION NUMBER: 10/023,617

; PRIOR FILING DATE: 2001-12-18

; PRIOR APPLICATION NUMBER: PCT/US01/49416

; PRIOR FILING DATE: 2001-12-18

; PRIOR APPLICATION NUMBER: 60/256,249

; PRIOR FILING DATE: 2000-12-18

; PRIOR APPLICATION NUMBER: 60/256,405

; PRIOR FILING DATE: 2000-12-18

; PRIOR APPLICATION NUMBER: 10/083,248

; PRIOR FILING DATE: 2001-10-22

; PRIOR APPLICATION NUMBER: PCT/US01/46717

; PRIOR FILING DATE: 2001-10-22

; PRIOR APPLICATION NUMBER: 60/242,324

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/242,518

; PRIOR FILING DATE: 2000-10-23

; PRIOR APPLICATION NUMBER: 60/241,989

; PRIOR FILING DATE: 2000-10-20

; NUMBER OF SEQ ID NOS: 86

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 29

; LENGTH: 376

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-176-306-29

Query Match 39.2%; Score 63.5; DB 14; Length 376;

Best Local Similarity 46.4%; Pred. No. 1.5;

Matches 13; Conservative 2; Mismatches 12; Indels 1; Gaps 1;

Qy 2 GHEQOHGLGHGKXKL-DDDLEHOGG 28

```

; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 33
; LENGTH: 476
; TYPE: PRT
; ORGANISM: mus musculus
US-10-360-849A-33

Query Match      39.2%; Score 63.5; DB 15; Length 476;
Best Local Similarity 40.6%; Pred. No. 1.9;
Matches 13; Conservative 0; Mismatches 14; Indels 5; Gaps 1;

Qy 1 HGHEQGHG-----LGHGKXKLLDDLEHGGH 27
Db 82 HAHSHDGHSEELHGHSHSHSHSLHGGH 113

RESULT 11
US-10-270-333-42
; Sequence 42, Application US/10270333
; Publication No. US20030092124A1
; GENERAL INFORMATION:
; APPLICANT: Craychik, Anibal
; TITLE OF INVENTION: ISOLATED G-PROTEIN COUPLED RECEPTORS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING GPCR PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF AS INSECTICIDAL TARGETS
; FILE REFERENCE: CL000733CON
; CURRENT APPLICATION NUMBER: US/10/270,333
; CURRENT FILING DATE: 2002-10-15
; PRIOR APPLICATION NUMBER: 60/168,677
; PRIOR FILING DATE: 1999-12-03
; PRIOR APPLICATION NUMBER: 60/175,691
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/191,638
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 198
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 42
; LENGTH: 385
; TYPE: PRT
; ORGANISM: Drosophila
US-10-270-333-42

Query Match      38.3%; Score 62; DB 14; Length 385;
Best Local Similarity 55.0%; Pred. No. 2.4;
Matches 11; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 1 HGHEQGHGHLGHGKXKLLDD 20
Db 217 HGHGHHGHGHHGHGKXKLLDD 236

RESULT 12
US-10-090-035-20
; Sequence 20, Application US/10090035
; Publication No. US20020170089A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
; TITLE OF INVENTION: Proteins and Uses Thereof
; FILE REFERENCE: 35718/242990
; CURRENT APPLICATION NUMBER: US/10/090,035
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/272,227
; PRIOR FILING DATE: 02/28/2001
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 92
; TYPE: PRT
; ORGANISM: Triticum aestivum
US-10-090-035-20

Query Match      37.7%; Score 61; DB 13; Length 92;

```

seq-b.rapb

Fri Mar 5 08:53:30 2004

Db 73 HGHTHDHDHGHSH-----EDLHGHSH 94

Best Local Similarity 36.7%; Pred. No. 0.75;
Matches 11; Conservative 6; Mismatches 9; Indels 4; Gaps 1;Qy 1 HGHQHQHGLGHG-----KKLDDDLHQQG 26
Db 48 HGHGHGHGSGHFEVRESLEEDINTGTG 77

RESULT 13
US-10-090-035-22
; Sequence 22, Application US/10050035
; Publication No. US20020170089A1
; GENERAL INFORMATION:
; APPLICANT: SIMMONS, Carl R.
; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
; TITLE OF INVENTION: Proteins and Uses Thereof
; FILE REFERENCE: 35718/242590
; CURRENT APPLICATION NUMBER: US/10/090,035
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/272,227
; PRIOR FILING DATE: 02/28/2001
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22
; LENGTH: 91
; TYPE: PRT
; ORGANISM: Triticum aestivum
US-10-090-035-22

Query Match 36.4%; Score 59; DB 13; Length 91;
Best Local Similarity 36.7%; Pred. No. 1.4;
Matches 11; Conservative 5; Mismatches 10; Indels 4; Gaps 1;Qy 1 HGHQHQHGLGHG-----KKLDDDLHQQG 26
Db 47 HGHGHGHGSGHFEVRESLEEDFNTRTG 76

RESULT 14
US-10-360-849A-36
; Sequence 36, Application US/10360849A
; Publication No. US20030220249A1
; GENERAL INFORMATION:
; APPLICANT: Discovery Genomics, Inc.
; APPLICANT: Hackett, Perry
; APPLICANT: Nasevicius, Aidan
; APPLICANT: Essner, Jeffrey
; APPLICANT: Clark, Karl
; APPLICANT: Larson, Jon
; APPLICANT: Ekker, Stephen
; APPLICANT: Roberg-Perez, Sharon
; APPLICANT: Madman, Shannon
; TITLE OF INVENTION: FACTORS FOR ANGIOGENESIS, VASCULOGENESIS, CARILAGE FORMATION,
; TITLE OF INVENTION: BONE FORMATION, AND METHODS OF USE THEREOF
; FILE REFERENCE: 3021.05US02
; CURRENT APPLICATION NUMBER: US/10/360,849A
; CURRENT FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: US 60/354,978
; PRIOR FILING DATE: 2002-02-07
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 36
; LENGTH: 469
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-360-849A-36

Query Match 36.1%; Score 58.5; DB 15; Length 469;
Best Local Similarity 40.7%; Pred. No. 8.7;
Matches 11; Conservative 1; Mismatches 10; Indels 5; Gaps 1;

Qy 1 HGHQHQHGLGHG-----KKLDDDLHQQG 27

Query Match 35.2%; Score 57; DB 14; Length 213;
Best Local Similarity 42.3%; Pred. No. 6;
Matches 11; Conservative 6; Mismatches 7; Indels 2; Gaps 1;Qy 2 GHEQHQHGLGHGKKLDDDLHQQGH 27
Db 10 GAVRQHGKANGYQVR--NDLEYWGAAH 33Search completed: March 4, 2004, 13:26:32
Job time : 42.6618 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2004, 13:02:33 ; Search time 17.2941 Seconds
(without alignments)
155.739 Million cell updates/sec

Title: SEQ-B

Perfect score: 162

Sequence: 1 HGHQHQHGLGHGKXKLLDDLEHGGHGV 28

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR_78.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	160	98.8	644	1 KGHUHI	kininogen, HMW pre
2	97.5	60.2	264	2 C25486	K-kininogen, HMW p
3	97.5	60.2	639	2 A25486	kininogen, HMW I p
4	91.5	56.5	290	2 C27115	K-kininogen, LMW p
5	91.5	56.5	315	2 A27115	major acute phase
6	85	52.5	619	1 KGBOH1	kininogen, HMW II
7	85	52.5	621	1 KGBOH1	kininogen, HMW I p
8	69	42.6	106	2 E87560	conserved hypothet
9	66	40.7	776	2 T02702	hypothetical prote
10	65.5	40.4	361	2 F87286	cation efflux fami
11	63.5	39.2	436	2 I49714	MHC H-2x/t-ws-link
12	63	38.9	735	2 T45059	hypothetical prote
13	62.5	38.6	535	2 S66148	gene pipsqueak pro
14	62.5	38.6	1085	2 S66149	gene pipsqueak pro
15	62	38.3	314	2 T35241	hypothetical prote
16	61.5	38.0	199	2 T48099	hypothetical prote
17	61	37.7	110	2 T07618	cold stress protel
18	61	37.7	2038	2 A43742	female sterile hom
19	60.5	37.3	232	2 E87520	hypothetical prote
20	60.5	37.3	398	2 T02681	probable zinc tran
21	60.5	37.3	515	2 T23089	hypothetical prote
22	60	37.0	191	2 D96701	unknown protein, 9
23	60	37.0	549	2 T15506	hypothetical prote
24	59.5	36.7	201	2 H82055	peptidyl-prolyl ci
25	59	36.4	335	2 D38532	hvpB protein - Rho
26	58.5	36.1	195	2 AD0024	peptidylprolyl iso
27	58.5	36.1	254	2 A31488	filaggrin - mouse
28	58.5	36.1	313	2 A28444	filaggrin precurs
29	58.5	36.1	670	2 F36791	hypothetical prote

ALIGNMENTS

RESULT 1

KGUHI

kininogen, HMW precursor [validated] - human

N;Alternate names: alpha-2-thiol proteinase inhibitor; preprokinogen; prokininogen
N;Contains: bradykinin (kallidin I); HMW kininogen I; HMW kininogen II; low molecular we
C;Species: Homo sapiens (man)

C;Date: 28-May-1986 #sequence revision 28-May-1986 #text change 08-Dec-2000

C;Accession: A01279; A25276; S32422; A91153; A24871; A27899; A27699; A31905; A34030; S024

R;Ohkubo, I.; Kurachi, K.; Takasawa, T.; Shiohara, H.; Sasaki, M.

Biochemistry 23, 5691-5697, 1984

A;Title: Isolation of a human cDNA for alpha-2-thiol proteinase inhibitor and its identit

A;Reference number: A90490; MUID:85122621; PMID:6441591

A;Accession: A01279

A;Molecule type: mRNA

A;Residues: 1-389 <CHK>

A;Cross-references: G3:K02566; NID:gl77889

R;Takagaki, Y.; Kitamura, N.; Nakanishi, S.

J. Biol. Chem. 260, 8601-8609, 1985

A;Title: Cloning and sequence analysis of cDNAs for human high molecular weight and low t

A;Reference number: A92544; MUID:85234582; PMID:2989293

A;Accession: A25276

A;Molecule type: mRNA

A;Residues: 1-592, 'I', 594-644 <TAK>

A;Cross-references: GB:M11437; NID:gl86751; PIDN:AAB59550.1; PID:9386852

R;Auerwald, E.A.; Roessler, D.; Mentele, R.; Asstalg-Machleidt, I.

FEBS Lett. 321, 93-97, 1993

A;Title: Cloning, expression and characterization of human kininogen domain 3.

A;Reference number: S32422; MUID:93223854; PMID:8467916

A;Accession: S32422

A;Molecule type: mRNA

A;Residues: 'ANSM', 253-377 <AUE>

A;Note: differences are due to known cloning artifacts

R;Lottspeich, F.; Kellermann, J.; Henschen, A.; Foerster, B.; Muller-Esterl, W.

Eur. J. Biochem. 152, 307-314, 1985

A;Title: The amino acid sequence of the light chain of human high-molecular-mass kininoge

A;Reference number: A91153; MUID:86030270; PMID:4054110

A;Accession: A91153

A;Molecule type: protein

A;Residues: 379-644 <LOT>

A;Note: the bradykinin sequence preceding the light chain sequence was not determined in

R;Kellermann, J.; Lottspeich, F.; Henschen, A.; Mueller-Esterl, W.

Eur. J. Biochem. 154, 471-478, 1986

A;Title: Completion of the primary structure of human high-molecular-mass kininogen. The

A;Reference number: A24871; MUID:86108361; PMID:3484703

A;Accession: A24871

A;Molecule type: protein

A;Residues: 'Z', 20-380 <KELI>

R;Kellermann, J.; Lottspeich, F.; Henschen, A.; Mueller-Esterl, W.

in Kinins IV, Greenbaum, L.M., and Margolius, H.S., ed., pp.85-89, Plenum Press, New Yor

A;Title: Amino acid sequence of the light chain of human high molecular mass kininogen.

A;Reference number: A27899

A;Accession: A27899

A:Molecule type: protein
A:Residues: 379-389, 'K', 390-407, 'Q', 409-644 <KEL2>
R:Minidroitin, T.; Carretero, O.A.; Proud, D.; Walz, D.; Scioli, A.G.
Biochem. Biophys. Res. Commun. 152, 519-526, 1988
A:Title: A new kinin moiety in human plasma kininogens.
A:Reference number: A27699; PMID:88209021; PMID:3365237
A:Accession: A27699
A:Molecule type: protein
A:Residues: 380-389 <MIN>
R:Maeda, H.; Matsumura, Y.; Kato, H.
J. Biol. Chem. 263, 16051-16054, 1988
A:Title: Purification and identification of [hydroxyprolyl(3)]bradykinin in ascitic fluid
A:Reference number: A31905; PMID:89034061; PMID:3182782
A:Accession: A31905
A:Molecule type: protein
A:Residues: 381-389 <MAE>
R:Saaguri, M.; Ikeda, M.; Ideishi, M.; Arakawa, K.
Biochem. Biophys. Res. Commun. 150, 511-516, 1988
A:Title: Identification of [hydroxyproline(3)]-lysyl-bradykinin released from human placenta
A:Reference number: A34030; PMID:88106632; PMID:3337729
A:Accession: A34030
A:Molecule type: protein
A:Residues: 380-389 <SAS>
R:Lenarcic, B.; Gabrijelcic, D.; Rozman, B.; Drobnic-Kosorok, M.; Turk, V.
Biol. Chem. Hoppe-Seyler 369, 257-263, 1988
A:Title: Human cathepsin B and cysteine proteinase inhibitors (CPIs) in inflammatory and non-inflammatory diseases
A:Reference number: S02482; PMID:89076517; PMID:3264507
A:Accession: S02482
A:Molecule type: protein
A:Residues: 1-19;189-192;310-314;381-389 <LENI>
R:Kato, H.; Matsumura, Y.; Maeda, H.
FEBS Lett. 232, 252-254, 1988
A:Title: Isolation and identification of hydroxyproline analogues of bradykinin in human plasma
A:Reference number: A61495; PMID:88211869; PMID:3366244
A:Accession: A61495
A:Molecule type: protein
A:Residues: 380-389 <KAT1>
A:Experimental source: urine
A:Note: this peptide had Pro-383 modified to 4-hydroxyproline
A:Accession: S61495
A:Molecule type: protein
A:Residues: 381-389 <KAT2>
A:Experimental source: urine
A:Note: this peptide had Pro-383 modified to 4-hydroxyproline
A:Accession: C61495
A:Molecule type: protein
A:Residues: 380-389 <KAT3>
R:Lenarcic, B.; Krasovec, M.; Ritonja, A.; Olafsson, I.; Turk, V.
FEBS Lett. 280, 211-215, 1991
A:Title: Inactivation of human cystatin C and kininogen by human cathepsin D.
A:Reference number: S14303; PMID:91192133; PMID:2013314
A:Accession: S1447
A:Molecule type: protein
A:Residues: 264-359, 'N', 361-375 <LEN2>
R:Little, S.S.; Johnson, D.A.
Biochem. J. 307, 341-346, 1995
A:Title: Human mast cell tryptase isoforms: separation and examination of substrate-specificity
A:Reference number: S55239; PMID:95251593; PMID:7733867
A:Accession: S55239
A:Molecule type: protein
A:Residues: 450-452, 'X', 454, 'X', 456 <LIT>
R:Straczek, J.; Maachi, F.; le Nguyen, D.; Becchi, M.; Heulin, M.H.; Nabet, P.; Belleville, J.
FEBS Lett. 373, 207-211, 1995
A:Title: Purification from human plasma of a tetrapeptide that potentiates insulin-like growth factor release
A:Reference number: S68059; PMID:96033974; PMID:7589467
A:Accession: S68059
A:Molecule type: protein
A:Residues: 431-434 <STR>
R:Kitamura, N.; Kitagawa, H.; Fukushima, D.; Takagaki, Y.; Miyata, T.; Nakanishi, S.
J. Biol. Chem. 260, 8610-8617, 1985
A:Title: Structural organization of the human kininogen gene and a model for its evolution
A:Reference number: A32545; PMID:85234583; PMID:12989294
A:Contents: annotation; gene organization

R.Pierce, J.V.

Ref. Proc. 27, 52-57, 1968

A:Title: Structural features of plasma kinins and kininogens.

A:Reference number: A91455; MUID:90255622; PMID:4952632

A:Accession: annotation; bradykinin

A:Contents: The HMW kininogen precursor and the LMW form are produced from the same gene t

C:Comment: Kininogen is a cysteine proteinase inhibitor, takes part in initiation of the

C:Comment: The glycine/histidine/lysine-rich region of HMW kininogen light chain is impos

C:Comment: The glycine/histidine/lysine-rich region of HMW kininogen light chain is impos

C:Comment: Bradykinin, released from kininogen by kallikrein, is a potent vasodilator, i

C:Comment: x-proline residue is present in the kininogen prior to the release of bradykinin.

C:Genetics:

A:Gene: GDB:KNG

A:Cross-references: GDB:125256; OMIM:228960

A:Map position: 3q27-3q27

A:Introns: 65/3; 102/3; 131/1; 188/3; 224/3; 253/1; 310/3; 346/3; 375/3

C:Superfamily: kininogen; cystatin homology

C:Keywords: alternative splicing; blood coagulation; cysteine proteinase inhibitor; dupli

F:1-18/Domain: signal sequence #status experimental <SIG>

F:19-644/Product: HMW kininogen I (prokininogen) #status experimental <MAT1>

F:19-379-390-644/Product: HMW kininogen II #status experimental <MAT2>

F:19-379/Domain: HMW kininogen heavy chain #status experimental <HC1>

F:19-131/Domain: cystatin homology <CY1>

F:142-253/Domain: cystatin homology <CY2>

F:264-375/Domain: cystatin homology <CY3>

F:380-389/Product: lysyl-bradykinin (kallidin II) #status experimental <KBDY>

F:381-389/Product: bradykinin (kallidin I) #status experimental <BDY>

F:390-644/Domain: HMW kininogen light chain #status experimental <LCH>

F:421-510/Region: glycine/histidine/lysine-rich 30-residue repeats

F:431-434/Product: low molecular weight growth promoting factor #status experimental <GPF>

F:19/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experimental

F:28-614,83-94,107-126,142-145,206-218,229-248,264-267,328-340,351-376/Disulfide bonds: 4

F:48/Binding site: carboxylate (Asn) (covalent) #status absent

F:169,205,294/Binding site: carboxylate (Asn) (covalent) #status experimental

F:379-380/Cleavage site: Met-lys (kallikrein) #status experimental

F:383/Modified site: 4-hydroxyproline (Pro) (partial) #status experimental

F:389-390/Cleavage site: Arg-Ser (kallikrein) #status experimental

F:401,533,542,546,557,571,533,628/Binding site: carboxylate (Thr) (covalent) #status ex

F:577/Binding site: carboxylate (Ser) (covalent) #status experimental

Query Match 98.8%; Score 160; DB 1; Length 644;

Best Local Similarity 96.4%; Pred.No. 3e-14; 1; Indels 0; Gaps 0;

Matches 27; Conservative 0; Mismatches 1;

QY 1 HGHEQHGLGHGHHKXKLDLDDLEHGGHV 28

DB 463 HGHEQHGLGHGHHKXKLDLDDLEHGGHV 490

RESULT 2

C25486

K-kininogen, HMW precursor - rat (fragment)

C:Species: Rattus norvegicus (Norway rat)

C:Date: 08-Mar-1989 #sequence_revision 08-Mar-1989 #text_change 30-Sep-1993

C:Accession: C25486

J. Biol. Chem. 262, 2190-2198, 1987

F:Kitagawa, H.; Kitamura, N.; Hayashida, H.; Miyata, T.; Nakanishi, S.

A:Reference number: A92625; MUID:87137443; PMID:3029068

A:Accession: C25486

A:Molecule type: DNA

A:Residues: 1-264 <KIT>

C:Comment: The nucleotide sequence was obtained from GenBank, release 55.0.

C:Superfamily: kininogen; cystatin homology

Query Match 60.2%; Score 97.5; DB 2; Length 264;

Best Local Similarity 50.0%; Pred.No. 4.7e-06; 9; Gaps 1;

Matches 19; Conservative 3; Mismatches 6; Indels 9;

QY 1 HGHEQQHGLGHGXKKLDD-----DLHQGGH 27
 |||:|||||:|||||
 dD 75 HGHQPHGLGHGHQKLDDLKQPEDGYDHRHPVG 110

RESULT 3
 A25486
 Amino acid sequence: Kininogen, HMW I precursor - rat
 N/Contains: bradykinin
 C/Species: Rattus norvegicus (Norway rat)
 C/Date: 08-Mar-1989 #sequence_revision 08-Mar-1989 #text_change 15-Nov-1996
 C/Accession: A25486
 R/Kitagawa, H.; Kitamura, N.; Hayashida, H.; Miyata, T.; Nakanishi, S.
 J. Biol. Chem. 262, 2190-2196, 1987
 A/Title: Differing expression patterns and evolution of the rat kininogen gene family.
 A/Reference number: A93625; MUID:87137443; PMID:3029068
 A/Accession: A25486
 A/Molecule type: mRNA
 A/Residues: 1-639 <KIT>
 A/Note: the authors translated the codon CAA for residue 347 as Asn
 C/Superfamily: kininogen; cystatin homology
 C/Keywords: alternative splicing
 F:1-18/Domain: signal sequence #status predicted <SIG>
 F:19-639/Product: kininogen, HMW I #status predicted <MAT>
 F:19-131/Domain: cystatin homology <CY1>
 F:142-253/Domain: cystatin homology <CY2>
 F:264-375/Domain: cystatin homology <CY3>
 Query Match 60.2%; Score 97.5; DB 2; Length 639;
 Best Local Similarity 50.0%; Pred. No. 1.2e-05;
 Matches 18; Conservative 3; Mismatches 6; Indels 9; Gaps 1;
 Qy 1 HGHEQQHGLGHGKXKLLDD-----DLEHQGGH 27
 Db 450 HGHQPHGLGHGHQKLLDLKQQRDGYDHRHPVGH 485
 RESULT 4
 C27115
 K-kininogen, LMW precursor - rat (fragments)
 C/Species: Rattus norvegicus (Norway rat)
 C/Date: 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change 20-Aug-1999
 C/Accession: C27115; A25488
 R/Fung, W.P.; Schreiber, G.
 J. Biol. Chem. 262, 9298-9308, 1987
 A/Title: Structure and expression of the genes for major acute phase alpha-1-protein
 A/Reference number: A92653; MUID:87250580; PMID:2439509
 A/Accession: C27115
 A/Molecule type: DNA
 A/Residues: 1-290 <FUN>
 R/Kageyama, R.; Kitamura, N.; Ohkubo, H.; Nakanishi, S.
 J. Biol. Chem. 262, 2345-2351, 1987
 A/Title: Differing utilization of homologous transcription initiation sites of rat K
 A/Reference number: A25488; MUID:87137465; PMID:3818598
 A/Accession: A25488
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-48 <KAG>
 C/Cross-references: GB:J02662; NID:G205071; PIDN:AAA41483.1; PID:G205072
 C/Superfamily: kininogen; cystatin homology
 F:19-65/Domain: cystatin homology (fragment) <CVS>
 Query Match 56.5%; Score 91.5; DB 2; Length 290;
 Best Local Similarity 47.2%; Pred. No. 3.5e-05;
 Matches 17; Conservative 4; Mismatches 6; Indels 9; Gaps 1;
 Qy 1 HGHEQQHGLGHGKXKLLDD-----DLEHQGGH 27
 Db 124 NGHQPHGLGHGHKXKLLDLKQQRDGYDHRHPVGH 159
 RESULT 5
 A27115
 major acute phase alpha-1 protein 1 - rat (fragments)
 C/Species: Rattus norvegicus (Norway rat)
 C/Date: 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change 16-Jul-1999
 C/Accession: A27115
 R/Fung, W.P.; Schreiber, G.

A>Title: Disulfide bonds in bovine HMW kininogens.
 A/Reference number: A94300
 A/Contents: annotation; disulfide bonds
 A/Note: article in Japanese
 C/Comment: The HMW kininogen precursor is produced from the same gene as the LMW form as
 C/Comment: Kininogen is a cysteine proteinase inhibitor, takes part in initiation of the
 C/Comment: The glycine/histidine/lysine-rich region of HMW kininogen light chain is impo
 C/Comment: Bradykinin, released from kininogen by kallikrein, is a potent vasodilator, i
 xyproline residue is present in the kininogen prior to the release of bradykinin.
 C/Superfamily: kininogen; cystatin homology
 C/Keywords: alternative splicing; blood coagulation; cysteine proteinase inhibitor; dupl
 F/1-18/Domain: signal sequence #status predicted <SIG>
 F/19-376/Product: HMW kininogen II #status predicted <MAT>
 F/19-130/Domain: HMW kininogen II heavy chain #status experimental <HCH>
 F/141-252/Domain: cystatin homology <CY1>
 F/141-252/Domain: cystatin homology <CY2>
 F/261-372/Domain: cystatin homology <CY3>
 F/377-386/Product: lysyl-bradykinin (kallidin II) #status experimental <KBDY>
 F/378-386/Product: bradykinin (kallidin I) #status experimental <BDY>
 F/387-619/Product: HMW kininogen II light chain #status experimental <LCH>
 F/418-488/Region: glycine/histidine/lysine-rich
 F/419-589/Region: pyrrolidone carboxylic acid (Gln) (in mature form) #status experimen
 F/27-589, 82-93, 106-125, 141-144, 205-217, 228-247, 261-264, 325-337, 348-367/Disulfide bonds:
 F/47/Binding site: carboxylate (Asn) (covalent) #status absent
 F/87, 168, 169, 204, 280/Binding site: carboxylate (Asn) (covalent) #status experimental
 F/136/Binding site: carboxylate (Thr) (covalent) (partial) #status experimental
 F/197/Binding site: carboxylate (Asn) (covalent) (partial) #status experimental
 F/376-377/Cleavage site: Met-Lys (kallikrein) #status experimental
 F/380/Modified site: 4-hydroxyproline (Pro) #status predicted
 F/386-387/Cleavage site: Arg-Ser (kallikrein) #status experimental
 F/396-400, 404, 510/Binding site: carboxylate (Ser) (covalent) #status experimental
 F/397, 398, 518, 522, 534, 546, 553, 568/Binding site: carboxylate (Thr) (covalent) #status ex
 F/496-497/Cleavage site: Arg-Thr (kallikrein) #status experimental

Query Match 52.5%; Score 85; DB 1; Length 619;
 Best Local Similarity 55.6%; Pred. No. 0.00061;
 Matches 15; Conservative 2; Mismatches 2; Indels 8; Gaps 1;
 Db 461 HGHQKHGLGHGKXKLDLDEHGGH 27
 |||:|||||
 461 HGHQKHGLGHGKXKLDLDEHGGH 479

RESULT 7
 KGBOR1
 N/Alternate names: alpha-2-thiol proteinase inhibitor; prokininogen
 N/Contains: bradykinin (kallidin); kininogen I; kininogen II; prokininogen
 C/Species: Bos primigenius taurus (cattle)
 C/Date: 14-Nov-1983 #sequence revision 14-Nov-1983 #text_change 22-Jun-1999
 C/Accession: A01281; A91923; A91938; A29559
 R/Kitamura, N.; Takagaki, Y.; Furuta, S.; Tanaka, T.; Nawa, H.; Nakanishi, S.
 Nature 305, 545-549, 1983
 A/Title: A single gene for bovine high molecular weight and low molecular weight kininog
 A/Reference number: A93317; MUID:84014106; PMID:6571699
 A/Accession: A01281
 A/Molecule type: mRNA
 A/Residues: 1-621 <KIT>
 A/Cross-references: GB:V01491; GB:K01757; NID:5491; PIDN:CAA24735.1; PID:5492
 R/Kato, H.; Nagasawa, S.; Suzuki, T.
 J. Biochem. 67, 313-323, 1970
 A/Title: Studies on the structure of bovine kininogen: cleavages of disulfide bonds and
 A/Reference number: A91923; MUID:70180420; PMID:4986212
 A/Accession: A91923
 A/Molecule type: protein
 A/Residues: 378-393 <KAT>
 R/Han, Y.N.; Komiya, M.; Iwanaga, S.; Suzuki, T.
 J. Biochem. 77, 55-68, 1975
 A/Title: Studies on the primary structure of bovine high-molecular-weight kininogen. Ami
 A/Reference number: A91938; MUID:75170265; PMID:1169237
 A/Accession: A91938
 A/Molecule type: protein
 A/Residues: 458-498 <HAN>

R/Sueyoshi, T.; Miyata, T.; Hashimoto, N.; Kato, H.; Hayashida, H.; Miyata, T.; Iwanaga,
 J. Biol. Chem. 262, 2768-2779, 1987
 A/Title: Bovine high molecular weight kininogen. The amino acid sequence, positions of c
 A/Reference number: A92627; MUID:87137530; PMID:3546295
 A/Accession: A29559
 A/Molecule type: protein
 A/Residues: 'Z', 20-123, 'I', 125-127, 'I', 129-378 <SUB>
 R/Iotepesich, F.; Kellermann, J.; Henschel, A.; Foertsch, B.; Muller-Esterl, W.
 Eur. J. Biochem. 152, 307-314, 1985
 A/Title: The amino acid sequence of the light chain of human high-molecular-mass kininog
 A/Reference number: A91153; MUID:86030270; PMID:4054110
 A/Contents: annotation; bovine cleavage sites; bovine carbohydrate binding sites
 R/Sueyoshi, T.; Miyata, T.; Kato, H.; Iwanaga, S.
 Seikagaku 56, 809, 1984
 A/Title: Disulfide bonds in bovine HMW kininogens.
 A/Reference number: A94300
 A/Contents: annotation; disulfide bonds
 A/Note: article in Japanese
 C/Comment: The HMW kininogen precursor is produced from the same gene as the LMW form as
 C/Comment: Kininogen is a cysteine proteinase inhibitor, takes part in initiation of the
 C/Comment: The glycine/histidine/lysine-rich region of HMW kininogen light chain is impo
 C/Comment: Bradykinin, released from kininogen by kallikrein, is a potent vasodilator, i
 xyproline residue is present in the kininogen prior to the release of bradykinin.
 C/Superfamily: kininogen; cystatin homology
 C/Keywords: alternative splicing; blood coagulation; cysteine proteinase inhibitor; dupl
 F/1-18/Domain: signal sequence #status predicted <SIG>
 F/19-621/Product: HMW prokininogen I #status predicted <MAT>
 F/19-379/Product: HMW kininogen I heavy chain #status experimental <HCH>
 F/19-130/Domain: cystatin homology <CY1>
 F/141-252/Domain: cystatin homology <CY2>
 F/263-374/Domain: cystatin homology <CY3>
 F/379-388/Product: lysyl-bradykinin (kallidin II) #status experimental <KBDY>
 F/380-388/Product: bradykinin (kallidin I) #status experimental <BDY>
 F/389-521/Product: HMW kininogen I light chain #status experimental <LCH>
 F/417-488/Region: glycine/histidine/lysine-rich
 F/127-591, 82-93, 106-125, 141-144, 205-217, 228-247, 263-266, 327-339, 350-369/Disulfide bonds:
 F/87, 168, 169, 204/Binding site: carboxylate (Asn) (covalent) #status experimental
 F/136/Binding site: carboxylate (Thr) (covalent) (partial) #status experimental
 F/197/Binding site: carboxylate (Asn) (covalent) (partial) #status experimental
 F/378-379/Cleavage site: Met-Lys (kallikrein) #status experimental
 F/382/Modified site: 4-hydroxyproline (Pro) #status predicted
 F/388-389/Cleavage site: Arg-Ser (kallikrein) #status experimental
 F/398, 406, 512/Binding site: carboxylate (Ser) (covalent) #status experimental
 F/399, 400, 520, 524, 536, 548, 553, 570/Binding site: carboxylate (Thr) (covalent) #status ex
 F/498-499/Cleavage site: Arg-Thr (kallikrein) #status experimental

Query Match 52.5%; Score 85; DB 1; Length 621;
 Best Local Similarity 55.6%; Pred. No. 0.00061;
 Matches 15; Conservative 2; Mismatches 2; Indels 8; Gaps 1;

QY 1 HGHQKHGLGHGKXKLDLDEHGGH 27
 |||:|||||
 Db 463 HGHQKHGLGHGKXKLDLDEHGGH 481

RESULT 8
 E87560

conserved hypothetical protein CC2510 [imported] - Caulobacter crescentus
 C/Species: Caulobacter crescentus
 C/Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001
 C/Accession: E87560
 R/Nierman, W.C.; Feldblym, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.J.
 R/Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
 n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
 Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
 A/Title: Complete Genome Sequence of Caulobacter crescentus.
 A/Reference number: A87249; MUID:21173698; PMID:11259647
 A/Accession: E87560
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-106 <STO>
 A/Cross-references: GB:AE005673; NID:g13424067; PIDN:AAK24481.1; GSPDB:GN00148

C;Species: Drosophila melanogaster
 C;Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 21-Jul-2000
 C;Accession: S66148
 R;Weber, U.; Siegel, V.; Mlodzik, M.
 EMO J. 14, 6247-6257, 1995
 A;Title: pipsqueak encodes a novel nuclear protein required downstream of seven-up for
 A;Reference number: S66148; MUID:96134923; PMID:8557044
 A;Accession: S66148
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-535 <WEB>
 A;Cross-references: EMBL:X90986; NID:g1149498; PIDN:CAA62473.1; PID:g1149499
 C;Genetics:
 A;Gene: pipsqueak
 C;Superfamily: POZ domain homology
 F;21-123/Domain: POZ domain homology <POZ>
 Query Match 38.6%; Score 62.5; DB 2; Length 535;
 Best Local Similarity 44.4%; Pred. No. 0.66;
 Matches 12; Conservative 0; Mismatches 4; Indels 11; Gaps 1;
 QY 1 HGHEQOQHGLGHGKXKLDLDDLEHQGGH 27
 DB 332 HEHEHNGHGHG-----GGH 347
 RESULT 14
 S66149
 gene pipsqueak protein A long form - fruit fly (Drosophila melanogaster)
 C;Species: Drosophila melanogaster
 C;Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 23-Sep-2002
 C;Accession: S66149; S66150; T45461
 R;Weber, U.; Siegel, V.; Mlodzik, M.
 EMO J. 14, 6247-6257, 1995
 A;Title: pipsqueak encodes a novel nuclear protein required downstream of seven-up for
 A;Reference number: S66148; MUID:96134923; PMID:8557044
 A;Accession: S66149
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-1085 <WEB>
 A;Cross-references: EMBL:X90986; NID:g1149498; PIDN:CAA62474.1; PID:g1149500
 A;Accession: S66150
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 'MQ', 428-1085 <WE2>
 A;Cross-references: EMBL:X90986; NID:g1149498; PIDN:CAA62475.1; PID:g1149501
 R;Horowitz, H.; Berg, C.A.
 Development 122, 1859-1871, 1996
 A;Title: The Drosophila pipsqueak gene encodes a nuclear BTB-domain-containing protein
 A;Reference number: 422972; MUID:96232300; PMID:8674425
 A;Accession: T45461
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: mRNA
 A;Residues: 1-355, 'E', 357-1005, 'H', 1021-1061, 'ERS', <HOR>
 A;Cross-references: EMBL:U48358; NID:g1203906; PIDN:AA47153.1; PID:g1203907
 A;Experimental source: tissue type ovarian
 C;Genetics:
 A;Gene: pipsqueak; psq
 A;Map position: II
 A;Introns: 427/3
 C;Function:
 A;Description: required for establishing polarity of the developing egg chamber
 C;Superfamily: BRCORE-2 protein; POZ domain homology
 F;21-123/Domain: POZ domain homology <POZ>
 Query Match 38.6%; Score 62.5; DB 2; Length 1085;
 Best Local Similarity 44.4%; Pred. No. 1.4;
 Matches 12; Conservative 0; Mismatches 4; Indels 11; Gaps 1;

QY 1 HGHEQOQHGLGHGKXKLDLDDLEHQGGH 27
 DB 332 HEHEHNGHGHG-----GGH 347

RESULT 15
 T35241
 hypothetical protein SC5C7.34 SC5C7.34 - Streptomyces coelicolor
 C;Species: Streptomyces coelicolor
 C;Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 05-Nov-1999
 C;Accession: T35241
 R;Seeger, K.J.; Harris, D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
 submitted to the EMBL Data Library, September 1998
 A;Reference number: Z21572
 A;Accession: T35241
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 1-314 <SEE>
 A;Cross-references: EMBL:AL031515; PIDN:CAA20646.1; GSPDB:GNC0070; SCOREDB:SC5C7.34
 A;Experimental source: strain A3(2)
 C;Genetics:
 A;Gene: SCOREDB:SC5C7.34

Query Match 38.3%; Score 62; DB 2; Length 314;
 Best Local Similarity 76.9%; Pred. No. 0.44;
 Matches 10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 HGHEQOQHGLGHGK 13
 DB 18 HGCGQGHGCHGH 30

Search completed: March 4, 2004, 13:11:22
 Job time : 18.2941 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2004, 12:53:12 ; Search time 11.1176 Seconds
(without alignments)
131.140 Million cell updates/sec

Title: SEQ-B
Perfect score: 162
Sequence: 1 HGHEQHQHGLGHGKXKXLDLLEHOGGHV 28

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	160	98.8	644	1 KNG HUMAN	P01042 homo sapien
2	97.5	60.2	639	1 KNG RAT	P08934 rattus norv
3	96.5	59.6	661	1 KNG MOUSE	O08677 mus musculus
4	85	52.5	619	1 KNG2 BOVIN	P01045 bos taurus
5	85	52.5	621	1 KNG1 BOVIN	P01044 bos taurus
6	71	43.8	693	1 CAUP DROME	P54268 drosophila
7	63.5	39.2	476	1 KE4 MOUSE	Q31125 mus musculus
8	62.5	38.6	449	1 CSUP DROME	Q9V3A4 drosophila
9	61	37.7	2038	1 FSH DROME	P13709 drosophila
10	60.5	37.3	515	1 KE4L CAEL	Q9XTQ7 caenorhabdi
11	59	36.4	118	1 SL03_RABIT	P50117 oryctolagus
12	59	36.4	335	1 HYPB_RHOCA	P26410 rhodobacter
13	58.5	36.1	469	1 KE4 HUMAN	Q92504 homo sapien
14	58.5	36.1	670	1 VG50 HSV11	Q00130 icotaurid h
15	57.5	35.5	336	1 FILA MOUSE	P11088 mus musculus
16	57	35.2	302	1 HYPB BRAJA	Q45257 bradyrhizob
17	57	35.2	337	1 ZNUA_HAEIN	P44526 haemophilus
18	56	34.6	439	1 COTL_YEAST	P32798 saccharomyc
19	55	34.0	352	1 KE4 BRARE	Q9PUB8 brachydanio
20	55	34.0	1061	1 TRC4_ECOLI	P27189 escherichia
21	55	34.0	1448	1 TRC5_ECOLI	P27190 escherichia
22	54	33.3	389	1 TPT3_ARATH	Q8LE59 arabidopsis
23	54	33.3	503	1 ZNTA_MOUSE	Q60738 mus musculus
24	54	33.3	507	1 ZNTL_RAT	Q62720 rattus norv
25	53.5	33.0	85	1 ANTF_SARPE	Q08617 sarcophaga
26	53.5	33.0	258	1 BOX5_NOTVI	P53771 notophthalm
27	52	32.1	382	1 PO3A_XENLA	P13365 xenopus lae
28	52	32.1	1174	1 PNLH_HUMAN	Q16825 homo sapien
29	51.5	31.8	212	1 SLVD_AERYH	Q07046 aeromonas h
30	51	31.5	306	1 CH38_DROME	P07183 drosophila
31	51	31.5	409	1 DP42_BACHD	Q9K9A8 bacillus ha
32	51	31.5	419	1 GSC DROME	P54366 drosophila
33	51	31.5	496	1 BAF1_KITMA	P33293 kluyveromyc

RESULT 1
KNG_HUMAN STANDARD; PRT; 644 AA.
AC P01042; P01043;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Kininogen precursor (Alpha-2-thiol proteinase inhibitor) [Contains: Bradykinin].
GN KNG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RS SEQUENCE FROM N.A. (ISOFORMS HMW AND LMW).
RC TISSUE=Liver;
RX MEDLINE=85234582; PubMed=2989293;
RA Takagaki Y., Kitamura N., Nakanishi S.;
RT "Cloning and sequence analysis of cDNAs for human high molecular weight and low molecular weight prekininogens. Primary structures of two human prekininogens.";
RT J. Biol. Chem. 260:8601-8609 (1985).
RN [2]
RS GENE STRUCTURE.
RX MEDLINE=85234583; PubMed=2989294;
RA Kitamura N., Kitagawa H., Fukushima D., Takagaki Y., Miyata T., Nakanishi S.;
RT "Structural organization of the human kininogen gene and a model for its evolution.";
RT J. Biol. Chem. 260:8610-8617 (1985).
RN [3]
RS SEQUENCE OF 1-401 FROM N.A.
RX MEDLINE=85122621; PubMed=6441591;
RA Ohkubo I., Kurachi K., Takasawa T., Shiokawa H., Sasaki M.;
RT "Isolation of a human cDNA for alpha 2-thiol proteinase inhibitor and its identity with low molecular weight kininogen.";
RT Biochemistry 23:5691-5697 (1984).
RN [4]
RS SEQUENCE OF 379-644.
RX MEDLINE=86030270; PubMed=4054110;
RA Lottspeich F., Kellermann J., Henschen A., Foeretsch B., Mueller-Esterl W.;
RT "The amino acid sequence of the light chain of human high-molecular-mass kininogen.";
RT Eur. J. Biochem. 152:307-314 (1985).
RN [5]
RS SEQUENCE OF 381-389.
RX MEDLINE=90255622; PubMed=4952632;
RA Pierce J.V.;
RT "Structural features of plasma kinins and kininogens.";
RT Fed. Proc. 27:52-57 (1968).
RN [6]
RS DISULFIDE BONDS.
RA Sueyoshi T., Miyata T., Kato H., Iwanaga S.;
RT "Disulfide bonds in bovine HMW kininogens.";

P32793 saccharomyc
P14734 drosophila
O89023 mus musculus
Q9eqv6 rattus norv
O91n81 drosophila
O17582 caenorhabdi
Q29175 sus scrofa
Q92qds thizobium m
P05425 enterococcu
Q48251 helicobacte
P21749 drosophila
P21750 drosophila

ALIGNMENTS

Fri Mar 5 08:53:31 2004

```

RL Seikagaku 56:809-808(1984).
RN [7]
RP CARBOHYDRATE-LINKAGE SITE ASN-294.
RX MEDLINE=22660472; PubMed=12754519;
RA Zhang H., Li X.-J., Martin D.B., Aebersold R.;
RT "Identification and quantification of N-linked glycoproteins using
RT hydrazide chemistry, stable isotope labeling and mass spectrometry.";
RL Nat. Biotechnol. 21:660-666(2003).
CC -!- FUNCTION: (1) Kininogens are inhibitors of thiol proteases; (2)
CC HMW-kininogen plays an important role in blood coagulation by
CC helping to position optimally prekallikrein and factor XI next to
CC factor XII. (3) HMW-kininogen inhibits the thrombin- and plasmin-
CC induced aggregation of thrombocytes; (4) the active peptide
CC bradykinin that is released from HMW-kininogen shows a variety of
CC physiological effects: (4A) influence in smooth muscle
CC contraction, (4B) induction of hypotension, (4C) natriuresis and
CC diuresis, (4D) decrease in blood glucose level, (4E) it is a
CC mediator of inflammation and causes (4E1) increase in vascular
CC permeability, (4E2) stimulation of nociceptors (4E3) release of
CC other mediators of inflammation (e.g. prostaglandins), (4F) it has
CC a cardioprotective effect (directly via bradykinin action,
CC indirectly via endothelium-derived relaxing factor action); (5)
CC LMW-kininogen inhibits the aggregation of thrombocytes; (6) LMW-
CC kininogen is in contrast to HMW-kininogen not involved in blood
CC clotting.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=HMW;
CC IsoId=P01042-1; Sequence=Displayed;
CC Name=LMW;
CC IsoId=P01042-2; Sequence=VSP_001261, VSP_001262;
CC -!- TISSUE SPECIFICITY: Plasma.
CC -!- PTM: Bradykinin is released from kininogen by plasma kallikrein.
CC -!- SIMILARITY: Contains 3 cystatin-like domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; K02566; AAB5497.1; -.
DR EMBL; M11437; AAB59550.1; -.
DR EMBL; M11438; AAB59550.1; JOINED.
DR EMBL; M11521; AAB59550.1; JOINED.
DR EMBL; M11522; AAB59550.1; JOINED.
DR EMBL; M11523; AAB59550.1; JOINED.
DR EMBL; M11524; AAB59550.1; JOINED.
DR EMBL; M11525; AAB59550.1; JOINED.
DR EMBL; M11526; AAB59550.1; JOINED.
DR EMBL; M11527; AAB59550.1; JOINED.
DR EMBL; M11528; AAB59550.1; JOINED.
DR EMBL; M11437; AAB59551.1; -.
DR EMBL; M11438; AAB59551.1; JOINED.
DR EMBL; M11521; AAB59551.1; JOINED.
DR EMBL; M11522; AAB59551.1; JOINED.
DR EMBL; M11523; AAB59551.1; JOINED.
DR EMBL; M11524; AAB59551.1; JOINED.
DR EMBL; M11525; AAB59551.1; JOINED.
DR EMBL; M11526; AAB59551.1; JOINED.
DR EMBL; M11527; AAB59551.1; JOINED.
DR EMBL; M11528; AAB59551.1; JOINED.
DR PIR; A01279; KGHU1.
DR PIR; A01280; KGHU1.
DR SWISS-2DPAGE; P01042; HUMAN.
DR Genew; HGNC:6383; KNG.
DR MIM; 228960; -.
DR GO; GO:0007596; P:blood coagulation; NAS.
DR GO; GO:0030146; P:diuresis; NAS.
DR GO; GO:0006954; P:inflammatory response; NAS.

```

GO; GO:0030147; P:natriuresis; NAS.
GO; GO:0006939; P:smooth muscle contraction; NAS.
InterPro; IPR000010; Cystatin.
InterPro; IPR002395; Kininogen.
Pfam; PF00031; cystatin; 3.
PRINTS; PR00334; KININOGEN.
SMART; SM00043; CY; 3.
PROSITE; PS00287; CYSTATIN; 2.
Glycoprotein; Plasma; Repeat; Thiol protease inhibitor; Vasodilator;
Bradykinin; Blood coagulation; Inflammatory response; Signal;
Alternative splicing; Pyrrolidone carboxylic acid.
SIGNAL 1 18
CHAIN 19 644 KININOGEN.
CHAIN 19 380 KININOGEN HEAVY CHAIN.
PEPTIDE 381 389 BRADYKININ.
CHAIN 390 644 KININOGEN LIGHT CHAIN.
DOMAIN 19 136 CYSTATIN-LIKE 1.
DOMAIN 137 258 CYSTATIN-LIKE 2.
DOMAIN 259 380 CYSTATIN-LIKE 3.
DOMAIN 420 510 HIS-RICH
(ASSOCIATED WITH CLOTTING ACTIVITY).
REPEAT 420 449
REPEAT 450 479
REPEAT 480 510
MOD RES 19 19
DISULFID 28 614
DISULFID 83 94
DISULFID 107 126
DISULFID 142 145
DISULFID 206 218
DISULFID 229 248
DISULFID 264 267
DISULFID 328 340
DISULFID 351 370
CARBOHYD 48 148
CARBOHYD 169 169
CARBOHYD 205 205
CARBOHYD 294 294
CARBOHYD 401 401
CARBOHYD 533 533
CARBOHYD 542 542
CARBOHYD 546 546
CARBOHYD 557 557
CARBOHYD 571 571
CARBOHYD 577 577
CARBOHYD 593 593
CARBOHYD 628 628
VARSPPLIC 402 427
N-LINKED (GLNAC. .) (POTENTIAL).
N-LINKED (GLNAC. .) (POTENTIAL).
N-LINKED (GLNAC. .) (POTENTIAL).
O-LINKED.
O-LINKED.
O-LINKED.
O-LINKED.
O-LINKED.
O-LINKED.
O-LINKED.
O-LINKED.
VSPPTSMAPAQDEERDSGKQGHTR -> SHIRCEYKGR
PKGAGPAREVS (in isoform LMW).
FTTd=VSP_001261.
Missing (in isoform LMW).
FTTd=VSP_001262.
T -> I (in REF. 1).
SEQUENCE 644 AA; 71,945 MW; 3132B4CBFA8FB7E CRC64;
Query Match 98.8%; Score 160; DB 1; Length 644;
Best Local Similarity 96.4%; Pred. No. 1.2e-14;
Matches 27; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 HGHEQHQGLGHGKXKLDLDDLEHQGHV 28
DB 463 HGHEQHQGLGHGKXKLDLDDLEHQGHV 490
|||||
RESULT 2
KNG RAT STANDARD; PRT; 639 AA.
ID KNG RAT
AC P08934; P08933;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 15-VAR-2004 (Rel. 43, Last annotation update)
DE Kininogen precursor [Contains: Bradykinin].
GN KNG.

OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OC NCBI_TaxID=10116;
 RN [1] _
 RP SEQUENCE FROM N.A. (ISOFORMS HMW AND LMW).
 RX MEDLINE=87137443; PubMed=3029068;
 RA Kitagawa H., Kitamura N., Hayashida H., Miyata T., Nakanishi S.;
 RA "Differing expression patterns and evolution of the rat kininogen
 gene family.";
 RL J. Biol. Chem. 262:2190-2198(1987).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM LMW).
 RX MEDLINE=86008264; PubMed=2413018;
 RA Furuto-Kato S., Matsumoto A., Kitamura N., Nakanishi S.;
 RA "Primary structures of the mRNAs encoding the rat precursors for
 bradykinin and T-kinin. Structural relationship of kininogens with
 major acute phase protein and alpha 1-cysteine proteinase
 inhibitor.";
 RL J. Biol. Chem. 260:12054-12059(1985).
 RN [3]
 RP SEQUENCE OF 1-65 FROM N.A.
 RC STRAIN=Buffalo;
 RX MEDLINE=87250580; PubMed=2439509;
 RA Fung W.-P., Schreiber G.;
 RA "Structure and expression of the genes for major acute phase alpha 1-
 protein (chitosanin) and kininogen in the rat.";
 RL J. Biol. Chem. 262:9298-9308(1987).
 RN [4]
 RP SEQUENCE OF 1-41 FROM N.A.
 RC STRAIN=Wistar; TISSUE=Liver;
 RX MEDLINE=87137465; PubMed=3818598;
 RA Kageyama R., Kitamura N., Ohkubo H., Nakanishi S.;
 RA "Differing utilization of homologous transcription initiation sites
 of rat K and T kininogen genes under inflammation condition.";
 RL J. Biol. Chem. 262:2345-2351(1987).
 CC -!- FUNCTION: (1) Kininogens are inhibitors of thiol proteases; (2)
 CC HMW-kininogen plays an important role in blood coagulation by
 CC helping to position optimally prekallikrein and factor XI next to
 CC factor XII; (3) HMW-kininogen inhibits the thrombin- and plasmin-
 CC induced aggregation of thrombocytes; (4) the active peptide
 CC bradykinin that is released from HMW-kininogen shows a variety of
 CC physiological effects: (4A) influence in smooth muscle
 CC contraction, (4B) induction of hypotension, (4C) natriuresis and
 CC diuresis, (4D) decrease in blood glucose level, (4E) it is a
 CC mediator of inflammation and causes (4E1) increase in vascular
 CC permeability, (4E2) stimulation of nociceptors (4E3) release of
 CC other mediators of inflammation (e.g. prostaglandins), (4F) it has
 CC a cardioprotective effect (directly via bradykinin action,
 CC indirectly via endothelium-derived relaxing factor action); (5)
 CC LMW-kininogen inhibits the aggregation of thrombocytes; (6) LMW-
 CC kininogen is in contrast to HMW-kininogen not involved in blood
 CC clotting.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=HMW;
 CC IsoId=P08934-1; Sequence=Displayed;
 CC Name=LMW;
 CC IsoId=P08934-2; Sequence=VSP_001265, VSP_001266;
 CC -!- TISSUE SPECIFICITY: Plasma.
 CC -!- PTM: Bradykinin is released from kininogen by plasma kallikrein.
 CC -!- MISCELLANEOUS: Rats express four types of kininogens: the
 CC classical HMW/LMW kininogens and two additional LMW-like
 CC kininogens: T-I and T-II.
 CC -!- SIMILARITY: Contains 3 cystatin-like domains.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

or send an email to license@isb-sib.ch.
 CC EMBL; L29428; AAA41486.1; -
 CC EMBL; M11884; AAA41487.1; -
 DR EMBL; M14369; AAA41484.1; -
 DR EMBL; M14369; AAA41485.1; ALT_SEQ.
 DR EMBL; M16455; AAA41482.1; -
 DR PIR; A25486; A25486.
 DR PIR; A28055; A28055.
 DR InterPro; IPR000010; Cystatin.
 DR InterPro; IPR002395; Kininogen.
 DR Pfam; PF00031; cystatin; 3.
 DR PRINTS; PRO0334; KININOGEN.
 DR SMART; SM00043; CY; 3.
 DR PROSITE; PS00287; CYSTATIN; 2.
 KW Glycoprotein; Plasma; Repeat; Thiol protease inhibitor; Vasodilator;
 KW Bradykinin; blood coagulation; Inflammatory response; signal;
 KW Alternative splicing; Multigene family.
 FT SIGNAL 1 18
 FT CHAIN 19 639 KININOGEN.
 FT CHAIN 19 380 KININOGEN HEAVY CHAIN.
 FT PEPTIDE 381 389 BRADYKININ.
 FT CHAIN 390 639 KININOGEN LIGHT CHAIN.
 FT DOMAIN 19 136 CYSTATIN-LIKE 1.
 FT DOMAIN 137 258 CYSTATIN-LIKE 2.
 FT DOMAIN 259 380 CYSTATIN-LIKE 3.
 FT DOMAIN 439 514 HIS-RICH.
 FT DISULFID 28 609 INTERCHAIN (BY SIMILARITY).
 FT DISULFID 83 94 BY SIMILARITY.
 FT DISULFID 107 126 BY SIMILARITY.
 FT DISULFID 142 145 BY SIMILARITY.
 FT DISULFID 206 218 BY SIMILARITY.
 FT DISULFID 229 248 BY SIMILARITY.
 FT DISULFID 264 267 BY SIMILARITY.
 FT DISULFID 328 340 BY SIMILARITY.
 FT DISULFID 351 370 BY SIMILARITY.
 FT CARBOHYD 82 82 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 127 127 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 169 169 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 205 205 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 294 294 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 529 529 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT VARSPLIC 402 433 VSPSYIARVQERDPEQGGPIHGHWLHAKQ -> RLINS
 FT CEYGRLLKAGAPAPQQAASVTVP (in isoform
 FT LMW).
 FT VARSPLIC 434 639 /FTid=VSP_001265.
 FT VARSPLIC 61 61 Missing (in isoform LMW).
 FT CONFLICT 61 61 /FTid=VSP_001266.
 FT CONFLICT 61 61 E -> K (IN REF. 2).
 SQ SEQUENCE 639 AA; 70933 MW; D3172DF94FF56AF5 CRC64;
 Query Match 60.2%; Score 97.5; DB 1; Length 639;
 Best Local Similarity 50.0%; Pred. No. 5.2e-06;
 Matches 18; Conservative 3; Mismatches 6; Indels 9; Gaps 1;
 QY 1 HGHEQHQGLGHEGKXKLDL-----DLEHQGH 27
 |||:|||||:|||||
 DB 450 HGHEQHQGLGHEGKXKLDLKKQREDGYDHRHPVGH 485
 RESULT 3
 ID KNG MOUSE
 AC KNG MOUSE STANDARD; PRT; 661 AA.
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Kininogen precursor [Contains: Bradykinin].
 GN KNG.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI_TaxID=10090;

EN SEQUENCE FROM N.A. (ISOFORMS HMW AND LMW).
 RP STRAIN=C57BL/6 X CBA; TISSUE=Liver;
 RC MEDLINE=97342556; PubMed=9199253;
 RX Takano M., Kondo J., Yamada K., Otani M., Sano K., Okamoto H.;
 RA "Molecular cloning of cDNAs for mouse low-molecular-weight and high-
 RT molecular-weight prekallikogens.";
 RL Biochim. Biophys. Acta 1352:222-230(1997).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM LMW).
 RC STRAIN=C57BL/6J; TISSUE=Placenta;
 RX MEDLINE=22354683; PubMed=12468851;
 OKazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
 RA Nikaido I., Oosato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
 RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gofjebori T.,
 RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
 RA Schramm L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
 RA Blake J.A., Bradt D., Brusci V., Chothia C., Corbani L.E., Cousins S.,
 RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
 RA Gaasterland T., Gariboldi M., Giasi C., Godzik A., Gough J.,
 RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
 RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
 RA Konoaga A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
 RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
 RA Nagashima T., Numata K., Okido T., Pavan W.J., Perte G., Pesole G.,
 RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
 RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
 RA Sandelin A., Schneider C., Sempile C.A., Setou M., Shimada K.,
 RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
 RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
 RA Wilming L.G., Wyshak-Boris A., Yanagisawa M., Yang L., Yang L.,
 RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
 RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
 RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
 RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
 RA Miyazaki A., Sakai K., Sakaki D., Shibata K., Shinagawa A.,
 RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
 RA Birney E., Hayashizaki Y.;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs.";
 RL Nature 420:563-573(2002).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORM LMW).
 RC TISSUE=Liver;
 RX MEDLINE=22368257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heide F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udgin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huylk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -1- FUNCTION: (1) Kininogens are inhibitors of thiol proteases; (2)
 CC HMW-kininogen plays an important role in blood coagulation by
 CC helping to position optimally prekallikrein and factor XI next to
 CC factor XII; (3) HMW-kininogen inhibits the thrombin- and plasmin-
 CC induced aggregation of thrombocytes; (4) the active peptide
 CC bradykinin that is released from HMW-kininogen shows a variety of
 CC physiological effects: (4A) influence in smooth muscle
 CC contraction, (4B) induction of hypotension, (4C) natriuresis and
 CC diuresis, (4D) decrease in blood glucose level, (4E) it is a
 CC mediator of inflammation and causes (4E1) increase in vascular
 CC permeability, (4E2) stimulation of nociceptors (4E3) release of
 CC other mediators of inflammation (e.g. prostaglandins), (4F) it has
 CC a cardioprotective effect (directly via bradykinin action), (5)
 CC indirectly via endothelium-derived relaxing factor action; (6) LMW-
 CC kininogen inhibits the aggregation of thrombocytes; (6) LMW-
 CC kininogen is in contrast to HMW-kininogen not involved in blood
 CC clotting (by similarity).
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=HMW;
 CC IsoId=O08677-1; Sequence=Displayed;
 CC Name=LMW;
 CC IsoId=O08677-2; Sequence=VSP_001263, VSP_001264;
 CC -1- TISSUE SPECIFICITY: Plasma
 CC -1- PM: Bradykinin is released from kininogen by plasma kallikrein.
 CC -1- SIMILARITY: Contains 3 cystatin-like domains.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: D84435; BAA19743.1; -;
 CC EMBL: D84435; BAA19742.1; -;
 CC EMBL: AK005547; BAB24115.1; -;
 CC EMBL: BC018158; AAI18158.1; -;
 CC MGD: MGI:1097705; Kog.
 CC InterPro: IPR000010; Cystatin.
 CC InterPro: IPR002395; Kininogen.
 CC Pfam: PF00031; cystatin; 3.
 CC PRINTS: PR00334; KININOGEN.
 CC SMART: SM00043; CY; 3.
 CC PROSITE: PS00287; CYSTATIN; 1.
 CC Glycoprotein; Plasma; Repeat; Thiol protease inhibitor; Vasodilator;
 CC Bradykinin; Blood coagulation; Inflammatory response; Signal;
 CC Alternative splicing.
 CC SIGNAL 1 18 POTENTIAL.
 CC CHAIN 19 661 KININOGEN.
 CC CHAIN 19 379 KININOGEN HEAVY CHAIN.
 CC CHAIN 380 388 BRADYKININ.
 CC CHAIN 389 661 KININOGEN LIGHT CHAIN.
 CC DOMAIN 19 135 CYSTATIN-LIKE 1.
 CC DOMAIN 136 257 CYSTATIN-LIKE 2.
 CC DOMAIN 258 379 CYSTATIN-LIKE 3.
 CC DOMAIN 439 524 HIS-RICH.
 CC DISULFID 28 631 INTERCHAIN (BY SIMILARITY).
 CC DISULFID 83 94 BY SIMILARITY.
 CC DISULFID 107 125 BY SIMILARITY.
 CC DISULFID 141 144 BY SIMILARITY.
 CC DISULFID 205 217 BY SIMILARITY.
 CC DISULFID 228 247 BY SIMILARITY.
 CC DISULFID 263 266 BY SIMILARITY.
 CC DISULFID 327 339 BY SIMILARITY.
 CC DISULFID 350 369 BY SIMILARITY.
 CC CARBOHYD 82 82 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC CARBOHYD 168 168 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC CARBOHYD 204 204 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC CARBOHYD 242 242 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC VARSPLIC 401 432 VSPPIAREQEDAEQTPHGHVLEKQ -> RLIRA
 CC CEYKGLSKAGAPAPQAESSQVKQ (in isoform
 CC LMW).
 CC /FTId=VSP_001263.
 CC Missing (in isoform LMW).
 CC /FTId=VSP_001264.
 CC SEQUENCE 661 AA; 73102 MW; 774460258D58796E CRC64;
 CC Query Match 59.6%; Score 96.5; DB 1; Length 661;

Best Local Similarity 50.0%; Pred. No. 7.5e-06;
Matches 18; Conservative 3; Mismatches 6; Indels 9; Gaps 1;

QY 1 HGHEQHQGLGHGHGXKLD-----DDLHQGGH 27
|||: |||||: |||
Db 460 HGHEQHQGLGHGHGXKLDLHQRHQRHGDHHTVGH 495

RESULT 4

KNH2_BOVIN
ID KMH2_BOVIN STANDARD; PRT; 619 AA.
AC P01045;
DT 21-JUL-1996 (Rel. 01, Created)
DT 21-JUL-1996 (Rel. 01, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Kininogen, HMW II precursor (Thiol proteinase inhibitor) [Contains:
DE Bradykinin]
OS Bos taurus [Bovine].
OC Bukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=84014106; PubMed=6571699;
RA Kitamura N., Takagaki Y., Furuto S., Tanaka T., Nawa H., Nakanishi S.;
RT "A single gene for bovine high molecular weight and low molecular
RT weight kininogens.";
RL Nature 305:545-549 (1993).
RN [2]
RP SEQUENCE OF 19-376.
RX MEDLINE=87137530; PubMed=3546295;
RA Sueyoshi T., Miyata T., Hashimoto N., Kato H., Hayashida H.,
RT Miyata T., Iwanaga S.;
RT "Bovine high molecular weight kininogen. The amino acid sequence,
RT positions of carbohydrate chains and disulfide bridges in the heavy
RT chain portion.";
RL J. Biol. Chem. 262:2768-2779 (1987).
RN [3]
RP SEQUENCE OF 376-391.
RX MEDLINE=70180420; PubMed=4986212;
RA Kato H., Nagasawa S., Suzuki T.;
RT "Studies on the structure of bovine kininogen: cleavages of disulfide
RT bonds and of methionyl bonds in kininogen-II.";
RL J. Biochem. 67:313-323 (1970).
RN [4]
RP SEQUENCE OF 387-455.
RX MEDLINE=76260155; PubMed=956151;
RA Han Y.N., Kato H., Iwanaga S., Suzuki T.;
RT "Primary structure of bovine plasma high-molecular-weight kininogen.
RT The amino acid sequence of a glycopeptide portion (fragment 1)
RT following the C-terminus of the bradykinin moiety.";
RL J. Biochem. 79:1201-1222 (1976).
RN [5]
RP SEQUENCE OF 456-496.
RX MEDLINE=75170265; PubMed=1169237;
RA Han Y.N., Komiya M., Iwanaga S., Suzuki T.;
RT "Studies on the primary structure of bovine high-molecular-weight
RT kininogen. Amino acid sequence of a fragment ('histidine-rich
RT peptide') released by plasma kallikrein.";
RL J. Biochem. 77:55-68 (1975).

CC -|- FUNCTION: (1) Kininogens are inhibitors of thiol proteases; (2)
CC helping to position optimally prekallikrein and factor XI next to
CC factor XII; (3) HMW-kininogen inhibits the thrombin- and plasmin-
CC induced aggregation of thrombocytes; (4) the active peptide
CC bradykinin that is released from HMW-kininogen shows a variety of
CC physiological effects: (4A) influence in smooth muscle
CC contraction, (4B) induction of hypotension, (4C) natriuresis and
CC diuresis, (4D) decrease in blood glucose level, (4E) it is a
CC mediator of inflammation and causes (4E1) increase in vascular
CC permeability, (4E2) stimulation of nociceptors (4E3) release of
CC other mediators of inflammation (e.g. prostaglandins), (4F) it has

CC a cardioprotective effect (directly via bradykinin action,
CC indirectly via endothelium-derived relaxing factor action).
CC -|- SUBCELLULAR LOCATION: Extracellular.
CC -|- ALTERNATIVE PRODUCTS:
CC Event-Alternative splicing; Named isoforms=2;
CC Name=HMW II;
CC IsoId=P01045-1; Sequence=Displayed;
CC Name=LMW II;
CC IsoId=P01047-1; Sequence=External;
CC TISSUE SPECIFICITY: Plasma.
CC -|- PTM: Bradykinin is released from kininogen by plasma kallikrein.
CC -|- SIMILARITY: Contains 3 cystatin-like domains.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC -----
CC EMBL; V01492; CAA24736.1; -.
DR PIR; A01282; KGBOH2.
DR HSP; P01038; I3A9.
DR InterPro; IPR000010; Cystatin.
DR InterPro; IPR002395; Kininogen.
DR Pfam; PF00031; cystatin; 3.
DR PRINTS; PRC0334; KININOGEN.
DR SMART; SM00043; CY; 3.
DR PROSITE; PS00287; CYSTATIN; 2.
KW Glycoprotein; Plasma; Repeat; Vasodilator; Alternative splicing;
KW Thiol protease inhibitor; Bradykinin; Blood coagulation; Signal;
KW Inflammatory response; Pyrrolidone carboxylic acid.
FT SIGNAL 1 18 KININOGEN, HMW II.
FT CHAIN 19 619 HEAVY CHAIN.
FT CHAIN 19 376 BRADYKININ.
FT PEPTIDE 378 386 LIGHT CHAIN.
FT CHAIN 387 619 CYSTATIN-LIKE 1.
FT DOMAIN 19 135 CYSTATIN-LIKE 2.
FT DOMAIN 136 256 CYSTATIN-LIKE 3.
FT DOMAIN 257 376 PYRROLIDONE CARBOXYLIC ACID.
FT MOD RES 19 19 N-LINKED (GLCNAC...).
FT CARBOHYD 87 87 O-LINKED (PARTIAL...).
FT CARBOHYD 136 136 N-LINKED (GLCNAC...). (OR 169).
FT CARBOHYD 168 168 N-LINKED (GLCNAC...). (PARTIAL).
FT CARBOHYD 197 197 N-LINKED (GLCNAC...).
FT CARBOHYD 204 204 N-LINKED (GLCNAC...).
FT CARBOHYD 280 280 O-LINKED.
FT CARBOHYD 400 400 INTERCHAIN.
FT DISULFID 27 589
FT DISULFID 82 93
FT DISULFID 106 125
FT DISULFID 141 144
FT DISULFID 205 217
FT DISULFID 228 247
FT DISULFID 261 264
FT DISULFID 325 337
FT DISULFID 348 367
FT VARIANT 398 398 T -> P.
FT VARIANT 401 401 L -> V.
FT VARIANT 454 454 H -> K.
SQ SEQUENCE 619 AA; 68710 MW; F04320A8EB0E0DA CRC64;

Query Match 52.8%; Score 85; DB 1; Length 619;
Best Local Similarity 55.6%; Pred. No. 0.00027;
Matches 15; Conservative 2; Mismatches 2; Indels 8; Gaps 1;

QY 1 HGHEQHQGLGHGHGXKLDLHQRHQRHGDHHTVGH 27
|||: |||||: |||

Db 461 HGHEQHQGLGHGHGXKLDLHQRHQRHGDHHTVGH 479

RESULT 5

KNLH_BOVIN
 ID KNLH_BOVIN STANDARD; PRT; 621 AA.
 AC P01044;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Kininogen, HMW I precursor (Thiol proteinase inhibitor) [Contains:
 DE Bradykinin].
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 ON NCBI_TaxID=9913;
 RX [1]
 RP SEQUENCE FROM N.A.
 RA Kitamura N., Takagaki Y., Furuto S., Tanaka T., Nawa H., Nakanishi S.;
 RA "A single gene for bovine high molecular weight and low molecular
 RT weight kininogens.";
 RL Nature 305:545-549 (1983).
 [2]
 RP SEQUENCE OF 19-378.
 RA MEDLINE=87137530; PubMed=3546295;
 RA Sueyoshi T., Miyata T., Hashimoto N., Kato H., Hayashida H.,
 RA Miyata T., Iwanaga S.;
 RT "Bovine high molecular weight kininogen. The amino acid sequence,
 RT positions of carboxylate chains and disulfide bridges in the heavy
 RT chain portion.";
 RL J. Biol. Chem. 262:2768-2779 (1987).
 [3]
 RP SEQUENCE OF 378-393.
 RA MEDLINE=70180420; PubMed=4966212;
 RA Kato H., Nagasawa S., Suzuki T.;
 RT "Studies on the structure of bovine kininogen: cleavages of disulfide
 RT bonds and of methionyl bonds in kininogen-II.";
 RL J. Biochem. 67:313-323 (1970).
 [4]
 RP SEQUENCE OF 458-498.
 RA MEDLINE=75170265; PubMed=1169237;
 RA Han Y.N., Komaya M., Iwanaga S., Suzuki T.;
 RT "Studies on the primary structure of bovine high-molecular-weight
 RT kininogen. Amino acid sequence of a fragment ('histidine-rich
 RT peptide') released by plasma kallikrein.";
 RL J. Biochem. 77:55-68 (1975).
 CC -!- FUNCTION: (1) Kininogens are inhibitors of thiol proteases; (2)
 CC HMW-kininogen plays an important role in blood coagulation by
 CC helping to position optimally prekallikrein and factor XI next to
 CC factor XII; (3) HMW-kininogen inhibits the thrombin- and plasmin-
 CC induced aggregation of thrombocytes; (4) the active peptide
 CC bradykinin that is released from HMW-kininogen shows a variety of
 CC physiological effects: (4A) influence in smooth muscle
 CC contraction, (4B) induction of hypotension, (4C) natriuresis and
 CC diuresis, (4D) decrease in blood glucose level, (4E) it is a
 CC mediator of inflammation and causes (4E1) increase in vascular
 CC permeability, (4E2) stimulation of nociceptors (4E3) release of
 CC other mediators of inflammation (e.g. prostaglandins), (4F) it has
 CC a cardioprotective effect (directly via bradykinin action,
 CC indirectly via endothelium-derived relaxing factor action).
 CC -!- SUBCELLULAR LOCATION: Extracellular.
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=HMW I;
 CC IsoId=P01044-1; Sequence=Displayed;
 CC Name=LMW I;
 CC IsoId=P01046-1; Sequence=External;
 CC -!- TISSUE SPECIFICITY: Plasma.
 CC -!- PTM: Bradykinin is released from kininogen by plasma kallikrein.
 CC -!- SIMILARITY: Contains 3 cystatin-like domains.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

EMBL; V01491; CAA24735.1; -
 PIR; A01281; KGSOHL.
 InterPro; IPR000010; Cystatin.
 InterPro; IPR002395; Kininogen.
 Pfam; PF00031; cystatin; 3.
 PRINTS; PR00334; KININOGEN.
 SMART; SM00043; CY; 3.
 PROSITE; PS00287; CYSTATIN; 2.
 Glycoprotein; Plasma; Repeat; Vascilator; Alternative splicing;
 Thiol protease inhibitor; Bradykinin; Blood coagulation;
 Inflammatory response; Signal; Pyrrolidone carboxylic acid.
 FT SIGNAL 1 18
 FT CHAIN 19 621 KININOGEN, HMW I.
 FT CHAIN 19 378 HEAVY CHAIN.
 FT PEPTIDE 380 388 BRADYKININ.
 FT CHAIN 389 621 LIGHT CHAIN.
 FT DOMAIN 19 135 CYSTATIN-LIKE 1.
 FT DOMAIN 136 257 CYSTATIN-LIKE 2.
 FT DOMAIN 258 378 CYSTATIN-LIKE 3.
 FT MOD_RES 19 19 PYRROLIDONE CARBOXYLIC ACID.
 FT CARBOHYD 87 87 N-LINKED (GLCNAC. . .).
 FT CARBOHYD 136 136 O-LINKED (PARTIAL).
 FT CARBOHYD 168 168 N-LINKED (GLCNAC. . .) (OR 169).
 FT CARBOHYD 197 197 N-LINKED (GLCNAC. . .) (PARTIAL).
 FT CARBOHYD 204 204 N-LINKED (GLCNAC. . .).
 FT DISULFID 27 591 INTERCHAIN.
 FT DISULFID 82 93
 FT DISULFID 106 125
 FT DISULFID 141 144
 FT DISULFID 205 217
 FT DISULFID 228 247
 FT DISULFID 263 266
 FT DISULFID 327 339
 FT DISULFID 350 369
 SQ SEQUENCE 621 AA; 68890 MW; D1680BFE3C55CD CRC64;

Query Match 52.5%; Score 85; DB 1; Length 621;
 Best Local Similarity 55.6%; Pred. No. 0.00027;
 Matches 15; Conservative 2; Mismatches 2; Indels 8; Gaps 1;

QY 1 HGHEQOQGLGEGHGXKLDLLEHQGH 27
 |||:|||||||
 Db 463 HGKQKQGLGEGHGXKLDLLEHQGH 481

RESULT 6
 CAUP_DROME STANDARD; PRT; 693 AA.
 AC P54269; Q9VJU0;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Homeobox protein caupolican.
 GN CAUP OR CG10605.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 ON NCBI_TaxID=7227;
 RX [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=96180722; PubMed=8620542;
 RA Gomez-Skarmeta J.-L., del Corral R.D., de la Calle-Mustienes E.,
 RA Ferrer-Marco D., Modolell J.;
 RT "Araucan and caupolican, two members of the novel ironquios complex,
 RL encode homeoproteins that control proneural and vein-forming genes.";
 RL Cell 85:95-110 (1996).
 [2]
 RP SEQUENCE FROM N.A.

RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McInerison D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,


```

[2]
RN SEQUENCE OF 45-82 FROM N.A.
RP STRAIN=New Zealand white;
RX MEDLINE=94198229; PubMed=8148323;
RA Mori S., Goto K., Goto F., Mitakami K., Ohkawara S., Yoshinaga M.;
RT "Dynamic changes in mRNA expression of neutrophils during the course
RT of acute inflammation in rabbits.";
RL Int. Immunol. 6:149-156(1994).
CC -!- SIMILARITY: Contains 2 EF-hand calcium-binding domains.
CC -!- SIMILARITY: Belongs to the S-100 family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF051849; AAC61771.1; -.
DR EMBL; D17404; BAA04227.1; -.
DR PIR; I46861; I46861.
DR HSSP; P80511; I88A.
DR InterPro; IPR001751; Csbp_S100.
DR InterPro; IPR002048; EF-hand.
DR Pfam; PF00036; ehand; 1.
DR Pfam; PF01023; S_100; 1.
DR ProDom; PD003407; Csbp_S100; 1.
DR ProDom; PD00012; EF-hand; 1.
DR ProSite; PS00018; EF_HAND; 1.
DR ProSite; PS00303; S100_CABP; 1.
DR Calcium-binding; Repeat.
KW NON_TER
FT 1 1 EF-HAND 1 (LOW AFFINITY) (POTENTIAL).
FT 22 22 EF-HAND 2 (HIGH AFFINITY) (POTENTIAL).
FT 53 64 2 X 8 AA TANDEN REPEATS OF G-H-G-H-
FT 103 118 G-H-S-H.
FT REPEAT 103 110 1.
FT REPEAT 111 118 2.
FT SEQUENCE 118 AA; 13292 MW; 7496118E21AD5C41 CRC64;
SQ
Query Match 36.4%; Score 59; DB 1; Length 118;
Best Local Similarity 69.2%; Pred. No. 0.18;
Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 HGEQHQGLGHGH 13
DB 104 HGHGSHGHGHGH 116
RESULT 12
HYPB RHOC
ID HYPB RHOC STANDARD; PRT; 335 AA.
AC P26410;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hydrogenase nickel incorporation protein hypb.
GN HYPB OR HUPB.
OS Rhodobacter capsulatus (Rhodospseudomonas capsulata).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodobacterales;
OC Rhodobacteraceae; Rhodobacter.
CX NCBI_TaxID=1061;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=ATCC 33303 / B10;
RX MEDLINE=93268090; PubMed=8497190;
RA Colbeau A., Richaud P., Toussaint B., Caballero F.J., Elster C.,
RA Delphin C., Smith R.L., Chabert J., Vignais P.M.;
RT "Organization of the genes necessary for hydrogenase expression in
RT Rhodobacter capsulatus. Sequence analysis and identification of two
RT hyp regulatory mutants.";
RL Mol. Microbiol. 8:15-29(1993).

```

```

RN SEQUENCE FROM N.A.
RP MEDLINE=91177833; PubMed=2007559;
RX Xu H.W., Wall J.D.;
RT "Clustering of genes necessary for hydrogen oxidation in Rhodobacter
RT capsulatus.";
RL J. Bacteriol. 173:2401-2405(1991).
CC -!- FUNCTION: Could be involved in nickel binding and accumulation.
CC -!- SIMILARITY: Belongs to the hypb/hupM family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X61007; CAA43326.1; -.
DR EMBL; M55089; AAA29246.1; -.
DR PIR; D38532; D38532.
DR InterPro; IPR004392; HypB.
DR InterPro; IPR002894; HypB_Ureg.
DR Pfam; PF01495; HypB_Ureg; 1.
DR TIGRfam; TIGR00073; hypB; 1.
DR Metal-binding; Nickel.
DR SEQUENCE 335 AA; 35342 MW; B7276C3E1BA0FD02 CRC64;
SQ
Query Match 36.4%; Score 59; DB 1; Length 335;
Best Local Similarity 40.7%; Pred. No. 0.56;
Matches 11; Conservative 1; Mismatches 13; Indels 2; Gaps 1;
QY 1 HGEQHQGLGHGHKXLDLLEHQGH 27
DB 64 HAHSHHAAGHG--AEADSDHPHAHG 89
RESULT 13
KE4 HUMAN
ID KE4 HUMAN STANDARD; PRT; 469 AA.
AC Q92504; Q9UIQ0;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Zinc transporter SLC39A7 (Solute carrier family 39 member 1)
DE (Histidine-rich membrane protein Ke4).
GN SLC39A7 OR HKE4 OR RING5.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
CX NCBI_TaxID=9606;
[1]
RN SEQUENCE FROM N.A.
RP TISSUE=Kidney;
RX MEDLINE=97001166; PubMed=8812499;
RA Ando A., Kikuti Y.Y., Shigenari A., Kawata H., Okamoto N., Shina T.,
RA Chen L., Ikemura T., Abe K., Kimura M., Inoko H.;
RT "cDNA cloning of the human homologues of the mouse Ke4 and Ke6 genes
RT at the centromeric end of the human MEC region.";
RL Genomics 35:600-602(1996).
[2]
RN SEQUENCE FROM N.A.
RP Vergara A., Lana I., Corella A., de Miguel C., Migliaccio M.,
RA Encio I.;
RT "Molecular cloning and characterization of the human KE4 gene and 5'
RT flanking region.";
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
[3]
RN SEQUENCE FROM N.A.
RP Tubby B.;
RT Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
[4]
RN SEQUENCE FROM N.A.
RP

```


TISSUE=Skin;
MEDLINE=22398257; PubMed=12477932;
A Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
A Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
A Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
A Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
A Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
A Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
A Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
A Rana S.S., Loggellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
A Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
A Richards S., Worley K.C., Hale S.J., Garcia A.M., Gay L.J., Hulyk S.W.,
A Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
A Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
A Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
A Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
A Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
A Butterfield Y.S.N., Krywinski M.I., Skalska U., Smailus D.E.,
A Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
A "Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences.";
Proc. Natl. Acad. Sci. U.S.A. 99:16999-16903 (2002).
[5]
P CHARACTERIZATION.
P PubMed=1452538;
P Taylor K.M., Morgan H.E., Johnson A., Nicholson R.I.;
P "Structure-function analysis of HKE4, a member of the new LIV-1
P subfamily of zinc transporters.";
P Biochem. J. 0:0(2003).
L -!- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
L -!- TISSUE SPECIFICITY: Major expression in placenta, lung, kidney
L and pancreas.
C -!- SIMILARITY: Belongs to the ZIP transporter (TC 2.A.5) family.
C KE4/Catsep subfamily.
C
C This SWISS-PROT entry is copyright. It is produced through a collaboration
C between the Swiss Institute of Bioinformatics and the EMBL outstation -
C the European Bioinformatics Institute. There are no restrictions on its
C use by non-profit institutions as long as its content is in no way
C modified and this statement is not removed. Usage by and for commercial
C entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
C or send an email to license@isb-sib.ch).
C
C EMBL: D82060; BAA11528.1; -;
C EMBL: AF117221; AAD12305.1; -;
C EMBL: AL031228; AAC02038.1; -;
C EMBL: BC000645; AAH00645.1; -;
C Genew: HGNC:14927; SLC39A7.
C MIM: 601416; -;
C GO: GO:0005624; C-membrane fraction; TAS.
C InterPro: IPR003689; Zn_transpt_Zip.
C Pfam: PF02535; Zip; 1.
C
C Transport; Transmembrane; Glycoprotein.
C TRANSMEM 10 30
C TRANSMEM 138 158 POTENTIAL.
C TRANSMEM 169 189 POTENTIAL.
C TRANSMEM 214 234 POTENTIAL.
C TRANSMEM 381 401 POTENTIAL.
C TRANSMEM 417 436 POTENTIAL.
C TRANSMEM 501 520 POTENTIAL.
C DOMAIN 30 114 HIS-RICH.
C DOMAIN 238 263 HIS-RICH.
C CARBOHYD 330 330 N-LINKED (GLCNAC... (POTENTIAL).
C A -> G (IN REF. 1 AND 2).
C E -> G (IN REF. 1 AND 2).
C S -> T (IN REF. 1 AND 2).
C CALUTEGGAVGEIAGAGGFWLPTTAGGFIYVATVSVLP
C ELLEFRASQSLLEVLGLGGLVIMVLAHLE -> VPFSL
C XEEQWTKLVQVQVLGSLGCHLLQVALST (IN REF. 1
C AND 2).
C
C GO: GO:0005624; C-membrane fraction; TAS.
C InterPro: IPR003689; Zn_transpt_Zip.
C Pfam: PF02535; Zip; 1.
C
C Transport; Transmembrane; Glycoprotein.
C TRANSMEM 10 30
C TRANSMEM 138 158 POTENTIAL.
C TRANSMEM 169 189 POTENTIAL.
C TRANSMEM 214 234 POTENTIAL.
C TRANSMEM 381 401 POTENTIAL.
C TRANSMEM 417 436 POTENTIAL.
C TRANSMEM 501 520 POTENTIAL.
C DOMAIN 30 114 HIS-RICH.
C DOMAIN 238 263 HIS-RICH.
C CARBOHYD 330 330 N-LINKED (GLCNAC... (POTENTIAL).
C A -> G (IN REF. 1 AND 2).
C E -> G (IN REF. 1 AND 2).
C S -> T (IN REF. 1 AND 2).
C CALUTEGGAVGEIAGAGGFWLPTTAGGFIYVATVSVLP
C ELLEFRASQSLLEVLGLGGLVIMVLAHLE -> VPFSL
C XEEQWTKLVQVQVLGSLGCHLLQVALST (IN REF. 1
C AND 2).
C
C GO: GO:0005624; C-membrane fraction; TAS.
C InterPro: IPR003689; Zn_transpt_Zip.
C Pfam: PF02535; Zip; 1.
C
C Transport; Transmembrane; Glycoprotein.
C TRANSMEM 10 30
C TRANSMEM 138 158 POTENTIAL.
C TRANSMEM 169 189 POTENTIAL.
C TRANSMEM 214 234 POTENTIAL.
C TRANSMEM 381 401 POTENTIAL.
C TRANSMEM 417 436 POTENTIAL.
C TRANSMEM 501 520 POTENTIAL.
C DOMAIN 30 114 HIS-RICH.
C DOMAIN 238 263 HIS-RICH.
C CARBOHYD 330 330 N-LINKED (GLCNAC... (POTENTIAL).
C A -> G (IN REF. 1 AND 2).
C E -> G (IN REF. 1 AND 2).
C S -> T (IN REF. 1 AND 2).
C CALUTEGGAVGEIAGAGGFWLPTTAGGFIYVATVSVLP
C ELLEFRASQSLLEVLGLGGLVIMVLAHLE -> VPFSL
C XEEQWTKLVQVQVLGSLGCHLLQVALST (IN REF. 1
C AND 2).
C
C GO: GO:0005624; C-membrane fraction; TAS.
C InterPro: IPR003689; Zn_transpt_Zip.
C Pfam: PF02535; Zip; 1.
C
C Transport; Transmembrane; Glycoprotein.
C TRANSMEM 10 30
C TRANSMEM 138 158 POTENTIAL.
C TRANSMEM 169 189 POTENTIAL.
C TRANSMEM 214 234 POTENTIAL.
C TRANSMEM 381 401 POTENTIAL.
C TRANSMEM 417 436 POTENTIAL.
C TRANSMEM 501 520 POTENTIAL.
C DOMAIN 30 114 HIS-RICH.
C DOMAIN 238 263 HIS-RICH.
C CARBOHYD 330 330 N-LINKED (GLCNAC... (POTENTIAL).
C A -> G (IN REF. 1 AND 2).
C E -> G (IN REF. 1 AND 2).
C S -> T (IN REF. 1 AND 2).
C CALUTEGGAVGEIAGAGGFWLPTTAGGFIYVATVSVLP
C ELLEFRASQSLLEVLGLGGLVIMVLAHLE -> VPFSL
C XEEQWTKLVQVQVLGSLGCHLLQVALST (IN REF. 1
C AND 2).
C
C GO: GO:0005624; C-membrane fraction; TAS.
C InterPro: IPR003689; Zn_transpt_Zip.
C Pfam: PF02535; Zip; 1.
C
C Transport; Transmembrane; Glycoprotein.
C TRANSMEM 10 30
C TRANSMEM 138 158 POTENTIAL.
C TRANSMEM 169 189 POTENTIAL.
C TRANSMEM 214 234 POTENTIAL.
C TRANSMEM 381 401 POTENTIAL.
C TRANSMEM 417 436 POTENTIAL.
C TRANSMEM 501 520 POTENTIAL.
C DOMAIN 30 114 HIS-RICH.
C DOMAIN 238 263 HIS-RICH.
C CARBOHYD 330 330 N-LINKED (GLCNAC... (POTENTIAL).
C A -> G (IN REF. 1 AND 2).
C E -> G (IN REF. 1 AND 2).
C S -> T (IN REF. 1 AND 2).
C CALUTEGGAVGEIAGAGGFWLPTTAGGFIYVATVSVLP
C ELLEFRASQSLLEVLGLGGLVIMVLAHLE -> VPFSL
C XEEQWTKLVQVQVLGSLGCHLLQVALST (IN REF. 1
C AND 2).
C
C GO: GO:0005624; C-membrane fraction; TAS.
C InterPro: IPR003689; Zn_transpt_Zip.
C Pfam: PF02535; Zip; 1.
C
C Transport; Transmembrane; Glycoprotein.
C TRANSMEM 10 30
C TRANSMEM 138 158 POTENTIAL.
C TRANSMEM 169 189 POTENTIAL.
C TRANSMEM 214 234 POTENTIAL.
C TRANSMEM 381 401 POTENTIAL.
C TRANSMEM 417 436 POTENTIAL.
C TRANSMEM 501 520 POTENTIAL.
C DOMAIN 30 114 HIS-RICH.
C DOMAIN 238 263 HIS-RICH.
C CARBOHYD 330 330 N-LINKED (GLCNAC... (POTENTIAL).
C A -> G (IN REF. 1 AND 2).
C E -> G (IN REF. 1 AND 2).
C S -> T (IN REF. 1 AND 2).
C CALUTEGGAVGEIAGAGGFWLPTTAGGFIYVATVSVLP
C ELLEFRASQSLLEVLGLGGLVIMVLAHLE -> VPFSL
C XEEQWTKLVQVQVLGSLGCHLLQVALST (IN REF. 1
C AND 2).
C
C GO: GO:0005624; C-membrane fraction; TAS.
C InterPro: IPR003689; Zn_transpt_Zip.
C Pfam: PF02535; Zip; 1.
C
C Transport; Transmembrane; Glycoprotein.
C TRANSMEM 10 30
C TRANSMEM 138 158 POTENTIAL.
C TRANSMEM 169 189 POTENTIAL.
C TRANSMEM 214 234 POTENTIAL.
C TRANSMEM 381 401 POTENTIAL.
C TRANSMEM 417 436 POTENTIAL.
C TRANSMEM 501 520 POTENTIAL.
C DOMAIN 30 114 HIS-RICH.
C DOMAIN 238 263 HIS-RICH.
C CARBOHYD 330 330 N-LINKED (GLCNAC... (POTENTIAL).
C A -> G (IN REF. 1 AND 2).
C E -> G (IN REF. 1 AND 2).
C S -> T (IN REF. 1 AND 2).
C CALUTEGGAVGEIAGAGGFWLPTTAGGFIYVATVSVLP
C ELLEFRASQSLLEVLGLGGLVIMVLAHLE -> VPFSL
C XEEQWTKLVQVQVLGSLGCHLLQVALST (IN REF. 1
C AND 2).
C
C GO: GO:0005624; C-membrane fraction; TAS.
C InterPro: IPR003689; Zn_transpt_Zip.
C Pfam: PF02535; Zip; 1.
C
C Transport; Transmembrane; Glycoprotein.
C TRANSMEM 10 30
C TRANSMEM 138 158 POTENTIAL.
C TRANSMEM 169 189 POTENTIAL.
C TRANSMEM 214 234 POTENTIAL.
C TRANSMEM 381 401 POTENTIAL.
C TRANSMEM 417 436 POTENTIAL.
C TRANSMEM 501 520 POTENTIAL.
C DOMAIN

```

Db      641 HGHGHGHGHGHGHGGR-----GPPGGH 662
|||||  || ||||| :
|||||  || ||||| :

RESULT 15
FILE MOUSE
ID -FILE MOUSE STANDARD; PRT; 336 AA.
AC P11088;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Filaggrin (Fragment).
GN FLG.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=8058903; PubMed=3680218;
RA Rothnagel J.A.; Mehrel T., Idler W.W., Roop D.R., Steinert P.M.;
RT "The gene for mouse epidermal filaggrin precursor. Its partial
RT characterization, expression, and sequence of a repeating filaggrin
RT unit."
RL J. Biol. Chem. 262:15643-15648 (1987).
RN [2]
RP REVISIONS.
RA Rothnagel J.A.;
RL Submitted (SEP-1988) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Aggregates keratin intermediate filaments and promotes
CC disulfide-bond formation among the intermediate filaments during
CC terminal differentiation of mammalian epidermis.
CC -!- PTM: FILAGRIN IS INITIALLY SYNTHESIZED AS A LARGE, INSOLUBLE,
CC HIGHLY PHOSPHORYLATED PRECURSOR CONTAINING MANY TANDEM COPIES
CC OF 248 AA. WHICH ARE NOT SEPARATED BY "LARGE LINKER". THE
CC PRECURSOR IS DEPOSITED AS KERATOHYALIN GRANULES. DURING TERMINAL
CC DIFFERENTIATION IT IS DEPHOSPHORYLATED & PROTEOLYTICALLY CLEAVED.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; J03458; AAM75559.1; -
CC FIR; A28444; A28444.
CC MGD; MGI:95553; Flg.
CC GO; GO:0001533; C:cornified envelope; IDA.
CC Phosphorylation; Developmental protein.
CC NON TER 1
CC SEQUENCE 336 AA; 35678 MW; 259F124D3AC0DB2D CRC64;

Query Match 35.5%; Score 57.5; DB 1; Length 336;
Best Local Similarity 32.5%; Pred. NO. 0.91;
Matches 13; Conservative 2; Mismatches 12; Indels 13; Gaps 1;

QY 1 HGHGHGHGHGHGH-----GHKKKLDLLEHQGGH 27
Db 170 HQHEQQRGHQHGHQHHEHQPSGHRQQSSGGRGHQH 209
|||||  ||||| :
|||||  ||||| :

Search completed: March 4, 2004, 13:08:06
Job time : 11.1176 secs

```

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2004, 13:02:03 ; Search time 52.7059 Seconds
(without alignments)
167.619 Million cell updates/sec

Title: SEQ-B

Perfect score: 162

Sequence: 1 HGHEQHQHGLGHGKXKLDLLEHQQGHV 28

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: SP_ARCHAEA:*
- 2: SP_BACTERIA:*
- 3: SP_FUNGI:*
- 4: SP_HUMAN:*
- 5: SP_INVERTEBRATE:*
- 6: SP_MAMMAL:*
- 7: SP_MHC:*
- 8: SP_ORGANELLE:*
- 9: SP_PHAGE:*
- 10: SP_PLANT:*
- 11: SP_RODENT:*
- 12: SP_VIRUS:*
- 13: SP_VERTEBRATE:*
- 14: SP_UNCLASSIFIED:*
- 15: SP_VIRUS:*
- 16: SP_BACTERIA:*
- 17: SP_ARCHAEA:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	97.5	60.2	126	11	O09016
2	72	44.4	292	16	O87XA8
3	72	44.4	574	16	Q828T3
4	71	43.8	325	5	Q8MR03
5	70.5	43.5	218	5	Q9V3P9
6	70	43.2	686	5	Q9VWS0
7	69	42.6	106	16	Q9ASE0
8	68.5	42.3	378	11	Q97KX1
9	68.5	42.3	378	11	Q80Y27
10	68.5	42.3	418	10	Q84NH0
11	66.5	41.0	206	5	Q9GTN1
12	66	40.7	279	10	Q93VS6
13	66	40.7	331	16	Q8CTV5
14	66	40.7	776	10	O81050
15	65.5	40.4	198	5	Q9NNY9
16	65.5	40.4	361	16	Q9ABC7

17	65	40.1	381	13	Q90WV0
18	64	39.5	99	10	Q9ZRC7
19	63.5	39.2	213	5	Q9GTN0
20	63.5	39.2	376	4	Q8TCH2
21	63.5	39.2	376	4	Q8NEW0
22	63	38.9	182	10	Q94DL7
23	63	38.9	211	16	Q8XHT2
24	63	38.9	340	16	Q8PFC3
25	63	38.9	421	10	Q947R8
26	63	38.9	735	5	Q9NES7
27	63	38.9	911	5	Q8GHA9
28	62.5	38.6	336	16	Q7URX6
29	62.5	38.6	1064	5	Q9V5N1
30	62.5	38.6	1085	5	Q24455
31	62	38.3	314	16	O86731
32	62	38.3	385	5	Q9VWX5
33	62	38.3	392	10	Q94B00
34	62	38.3	450	5	Q27920
35	62	38.3	495	5	Q8SZF0
36	61.5	38.0	199	10	Q9LYE2
37	61	37.7	110	10	O64396
38	61	37.7	198	2	Q8RPX9
39	61	37.7	245	10	Q9XEL3
40	61	37.7	293	16	Q816T1
41	61	37.7	1110	5	Q8IRN6
42	61	37.7	2038	5	Q9W3L3
43	60.5	37.3	232	16	Q9A6A4
44	60.5	37.3	398	10	O81036
45	60	37.0	79	10	Q9W435

ALIGNMENTS

RESULT 1

O09016	PRELIMINARY;	PRT;	126 AA.
AC	O09016;		
DT	01-JUL-1997 (TrEMBLrel. 04, Created)		
DT	01-JUL-1997 (TrEMBLrel. 04, Last sequence update)		
DT	01-OCT-2002 (TrEMBLrel. 22, Last annotation update)		
DE	K-kininogen (Fragment).		
GN	KNGK.		
OS	Rattus norvegicus (Rat).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.		
OX	NCBI_TaxID=10116;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=Donryu;		
RX	MEDLINE=37468288; PubMed=9321484;		
RA	Harris E.L., Grigor M.R., Innes B.A., Harrap S.B., Koike G.,		
RA	Jacob H.J.;		
RT	"Strain-specific deletions in exon 10 of rat K-kininogen and T1-		
RT	kininogen genes allow mapping of both genes to rat chromosome 11."		
RL	Mamm. Genome 8:791-792(1997).		
DR	EMBL; AF003623; AAC09070.1; -.		
DR	InterPro; IPR002395; Kininogen.		
DR	PRINTS; PR00334; KININOGEN.		
FT	NON TER	1	
FT	VARIANT	24	24 D -> N.
FT	VARIANT	87	89 ROK -> HGQ.
FT	VARIANT	115	115 S -> N.
FT	NON TER	126	126
SQ	SEQUENCE	126 AA;	14092 MW; 9CCDF8751DA49C88 CRC64;

Query Match	60.28;	Score	97.5;	DB	11;	Length	126;
Best local similarity	50.04;	Pred. No.	1.8e-06;				
Matches	18;	Conservative	3;	Mismatches	6;	Indels	9;
Gaps	1;						
QY	1	HGHEQHQHGLGHGKXKLDLLEHQQGHV	27				
DB	38	HGHEQHQHGLGHGKXKLDLLEHQQGHV	73				

RESULT 2

Q87XA8 ID Q87XA8 PRELIMINARY; PRT; 292 AA.
 AC Q87XA8;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Membrane protein, putative.
 GN PSP704280.
 OS Pseudomonas syringae (pv. tomato).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Pseudomonas.
 OX NCBI_TaxID=323;
 RN [1]
 SEQUENCE FROM N.A.
 RP STRAIN=DC3000;
 RC Buell R., Joardar V., Khouri H., Fedorova N., Tran B., Russell D.,
 RA Berry K., Utterback T., Van Aken S., Feldblyum T., Gwinn M.,
 RA Dodson R., DeBoy R., Durkin A., Kolonay J., Madupu R., Daugherty S.,
 RA Brinkac L., Bearan M., Haft D., Selengut J., Nelson W., Davidsen T.,
 RA White O., Fraser C., Collier A.;
 RT "Complete sequence of Pseudomonas syringae";
 RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AE016871; AAC57731.1; -.
 DR TIGR; PSPT04280; -.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0015099; F:nickel ion transporter activity; IEA.
 DR GO; GO:0015675; P:nickel ion transport; IEA.
 DR InterPro; IPR004688; NiCo.
 DR Pfam; PF03824; NiCo; 1.
 DR Complete proteome.
 KW Complete proteome.
 SQ SEQUENCE 292 AA; 31560 MW; B456C00B314B5ED7 CRC64;

Query Match 44.4%; Score 72; DB 16; Length 292;
 Best Local Similarity 48.1%; Pred. No. 0.024;
 Matches 13; Conservative 2; Mismatches 8; Indels 4; Gaps 1;

QY 1 HGHQHQHGLGHGKXKLDLLEHGGH 27
 DB 123 HGHDETHSHNH-----DHDLEHHGH 150

RESULT 3

Q828T3 ID Q828T3 PRELIMINARY; PRT; 574 AA.
 AC Q828T3;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Hypothetical protein.
 GN SAV6579.
 OS Streptomyces avermitilis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Streptomycineae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=33903;
 RN [1]
 SEQUENCE FROM N.A.
 RP STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
 RC MEDLINE=21477403; PubMed=11572948;
 RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
 RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osone T.,
 RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
 RT "Genome sequence of an industrial microorganism Streptomyces
 avermitilis: deducing the ability of producing secondary
 metabolites.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
 RX MEDLINE=22608305; PubMed=12692562;
 RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,

RA Sakaki Y., Hattori M., Omura S.;
 RT "Complete genome sequence and comparative analysis of the industrial
 RT microorganism Streptomyces avermitilis.";
 RL Nat. Biotechnol. 21:526-531(2003).
 DR EMBL; AP005047; BAC74290.1; -.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:001584; P:rhodopsin-like receptor activity; IEA.
 DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. .; IEA.
 DR InterPro; IPR000276; GPCR Rhodpsn.
 DR InterPro; IPR000437; Prok_lipoprot_S.
 DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
 DR PROSITE; PS00013; PROKAR LIPOPROTEIN; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 574 AA; 59898 MW; 8061208E633BEE9E CRC64;

Query Match 44.4%; Score 72; DB 16; Length 574;
 Best Local Similarity 52.2%; Pred. No. 0.05; 9; Indels 0; Gaps 0;
 Matches 12; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

QY 1 HGHQHQHGLGHGKXKLDLLEH 23
 DB 418 HKENENHGHGHGKHDDHDKH 440

RESULT 4

Q8MR03 ID Q8MR03 PRELIMINARY; PRT; 325 AA.
 AC Q8MR03;
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE LP03275P (Fragment).
 GN CAUP OR CGI0605.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 SEQUENCE FROM N.A.
 RP STRAIN=Berkley;
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
 RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
 RA George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
 RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
 RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
 RA Celniker S.;
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY122206; AAM52718.1; -.
 DR FlyBase; FBgn0015919; caup.
 DR GO; GO:0005634; C:nucleus; IEA.
 DR GO; GO:0003677; F:DNA binding; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR InterPro; IPR003893; Iroquois_homeo.
 DR SMART; SM00548; IRO; 1.
 FT NON TER
 SQ SEQUENCE 325 AA; 35014 MW; 07AC5FF3D4322FCD CRC64;

Query Match 43.8%; Score 71; DB 5; Length 325;
 Best Local Similarity 50.0%; Pred. No. 0.037;
 Matches 14; Conservative 1; Mismatches 7; Indels 6; Gaps 1;

QY 1 HGHQHQHGLGHGKXKLDLLEHGGH 28
 DB 288 HGHGHGHGHLGHG-----GLGHGHGHM 309

RESULT 5

Q9V3P9 ID Q9V3P9 PRELIMINARY; PRT; 218 AA.
 AC Q9V3P9;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE BG:DS062238.4 protein.
 OS BG:DS062238.4 OR C53474.
 GN Drosophila melanogaster (Fruit fly).
 OS Rukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_taxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkley;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foaier C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA "The genome sequence of Drosophila melanogaster."
 RL Science 287:2185-2195(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkley;
 RX MEDLINE=99403001; PubMed=10471707;
 RA Ashburner M., Mitra S., Roote J., Lewis S.E., Blazej R.G., Davis T.,
 RA Doyle C., Galle R., George R., Harris N., Hartzell G., Harvey D.,
 RA Hong L., Houston K., Hoskins R., Johnson G., Martin C., Moshrefi A.,
 RA Palazzolo M., Reese M.G., Spradling A., Teang G., Wan K., Whitelaw K.,
 RA Celnik S., Rubin G.M.;
 RA "An exploration of the sequence of a 2.9-Mb region of the genome of
 RT Drosophila melanogaster: the Adh region.";
 RL Genetics 153:179-219(1999).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkley;
 RA Celnik S.E., Agbayani A., Arcaina T.T., Baxter E., Blazej R.G.,
 RA Butenhoff C., Champe M., Chavez C., Chew M., Ciesiolka L., Doyle C.M.,
 RA Farfan D.E., Galle R., George R.A., Harris N.L., Hoskins R.A.,
 RA Houston K.A., Humasti S.R., Karra K., Kearney L., Kim E., Lee B.,
 RA Lewis S., Li P., Lonatan M.A., Mazda P., Moshrefi A.R., Moshrefi M.,
 RA Nixon K., Pacleb J.M., Park S., Pfeiffer B., Poon L., Siqueira A.,
 RA Sethi H., Snir E., Svirskas R.R., Wan K.H., Weinburg T., Zhang R.,
 RA Zieran L.L., Rubin G.M.;

Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 ENMBL: AE003643; AAF53397.1; -.
 DR ENMBL: AE003409; AAF44875.1; -.
 DR FlyBase: FBgn0028871; BG:DS062238.4.
 DR InterPro: IPR000618; Insect cuticle.
 DR Pfam: PF00379; Chitin_bind_4; 1.
 DR PRINTS: PR00947; CUTICLE.
 DR PROSITE: PS00233; CUTICLE; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 218 AA; 24371 NW; 8C920ECC351E3529 CRC64;
 Query Match 43.5%; Score 70.5; DB 5; Length 218;
 Best Local Similarity 38.9%; Pred. No. 0.028; 10; Indels 9; Gaps 1;
 Matches 14; Conservative 3; Mismatches 10; Indels 9; Gaps 1;
 1 HQF-----EQQHGGLGHGKXKLDLLEHQGGH 27
 182 HHGSSSHSYSLKQEHGHHGSHGQHGFGHGY 217
 RESULT 6
 Q9VWS0 PRELIMINARY; PRT; 686 AA.
 ID Q9VWS0
 AC Q9VWS0
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE CG6632 protein.
 GN CG6632
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkley;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foaier C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA "The genome sequence of Drosophila melanogaster."
 RL Science 287:2185-2195(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkley;
 RX MEDLINE=99403001; PubMed=10471707;
 RA Ashburner M., Mitra S., Roote J., Lewis S.E., Blazej R.G., Davis T.,
 RA Doyle C., Galle R., George R., Harris N., Hartzell G., Harvey D.,
 RA Hong L., Houston K., Hoskins R., Johnson G., Martin C., Moshrefi A.,
 RA Palazzolo M., Reese M.G., Spradling A., Teang G., Wan K., Whitelaw K.,
 RA Celnik S., Rubin G.M.;
 RA "An exploration of the sequence of a 2.9-Mb region of the genome of
 RT Drosophila melanogaster: the Adh region.";
 RL Genetics 153:179-219(1999).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkley;
 RA Celnik S.E., Agbayani A., Arcaina T.T., Baxter E., Blazej R.G.,
 RA Butenhoff C., Champe M., Chavez C., Chew M., Ciesiolka L., Doyle C.M.,
 RA Farfan D.E., Galle R., George R.A., Harris N.L., Hoskins R.A.,
 RA Houston K.A., Humasti S.R., Karra K., Kearney L., Kim E., Lee B.,
 RA Lewis S., Li P., Lonatan M.A., Mazda P., Moshrefi A.R., Moshrefi M.,
 RA Nixon K., Pacleb J.M., Park S., Pfeiffer B., Poon L., Siqueira A.,
 RA Sethi H., Snir E., Svirskas R.R., Wan K.H., Weinburg T., Zhang R.,
 RA Zieran L.L., Rubin G.M.;


```

DR Pfam: PF01545; Cation efflux; 1.
DR TIGRFAMs: TIGR01297; CDF; 1.
KW Hypothetical protein.
SQ SEQUENCE 378 AA; 41790 MW; B98AC19C9A045006 CRC64;

Query Match      42.3%; Score 68.5; DB 11; Length 378;
Best Local Similarity 48.1%; Pred. No. 0.1;
Matches 13; Conservative 2; Mismatches 11; Indels 1; Gaps 1;

QY 2 GHEQQHGLGHGKXKL-DDDLHQGGH 27
DQ 163 GHGSHGSGHGSHSLFNGALDHSGH 189

RESULT 9
Q80Y27 PRELIMINARY; PRT; 378 AA.
ID Q80Y27
AC Q80Y27
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Zinc transporter like 2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Embryo;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins P.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore I., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raba S.S., Lequellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalls D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RL and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Embryo;
RA Strausberg R.;
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC050193; AAH50193.1; -.
DR GO: GO:0016020; C:membrane; IEA.
DR GO: GO:0008324; F:cation transporter activity; IEA.
DR GO: GO:0006812; P:cation transport; IEA.
DR InterPro: IPR002524; Cation efflux.
DR Pfam: PF01545; Cation efflux; 1.
DR TIGRFAMs: TIGR01297; CDF; 1.
SQ SEQUENCE 378 AA; 41776 MW; B989619A9A045000 CRC64;

Query Match      42.3%; Score 68.5; DB 11; Length 378;
Best Local Similarity 48.1%; Pred. No. 0.1;
Matches 13; Conservative 2; Mismatches 11; Indels 1; Gaps 1;

QY 2 GHEQQHGLGHGKXKL-DDDLHQGGH 27
DQ 163 GHGSHGSGHGSHSLFNGALDHSGH 189

```

```

RESULT 10
Q84NH0 PRELIMINARY; PRT; 418 AA.
ID Q84NH0
AC Q84NH0
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Zinc transporter.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzoideae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Jiu Gaigang; TISSUE=Root;
RA Huang J., Zhang H., Hou F.;
RT "Cloning and characterization of a ZAT-like zinc transporter cDNA from
RT rice.";
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY266290; AAP31024.1; -.
DR GO: GO:0016020; C:membrane; IEA.
DR GO: GO:0008324; F:cation transporter activity; IEA.
DR GO: GO:0006812; P:cation transport; IEA.
DR InterPro: IPR002524; Cation efflux.
DR Pfam: PF01545; Cation efflux; 1.
DR TIGRFAMs: TIGR01297; CDF; 1.
SQ SEQUENCE 418 AA; 45547 MW; EA197CFB61A5C7E5 CRC64;

Query Match      42.3%; Score 68.5; DB 10; Length 418;
Best Local Similarity 41.0%; Pred. No. 0.11;
Matches 16; Conservative 1; Mismatches 9; Indels 13; Gaps 2;

QY 1 EGHEQQHGLGHGKXKL-DDDLHQGGH 28
DQ 184 HGHEGSHGSHGSHSLFNGALDHSGH 220

RESULT 11
Q9GTN1 PRELIMINARY; PRT; 206 AA.
ID Q9GTN1
AC Q9GTN1
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE DS06238.4-like protein (fragment).
GN EG:DS06238.4.
OS Drosophila simulans (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7240;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=North Carolina 17;
RA Schmid K.J., Aquadro C.F.;
RT "The evolutionary analysis of 'orphans' from the Drosophila genome
RT identifies incorrectly annotated and rapidly evolving genes.";
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF264919; AAG10258.1; -.
DR FlyBase: FBgn0042861; Dsim\BG:DS06238.4.
DR InterPro: IPR000618; Insect cuticle.
DR Pfam: PF00379; Chitin bind_4; 1.
DR PRINTS: PR00947; CUTICLE.
DR PROSITE: PS00233; CUTICLE; 1.
FT NON TER 206
SQ SEQUENCE 206 AA; 23011 MW; E24B8C189BC0746D CRC64;

Query Match      41.0%; Score 66.5; DB 5; Length 206;
Best Local Similarity 45.7%; Pred. No. 0.1;
Matches 14; Conservative 0; Mismatches 13; Indels 3; Gaps 1;

```

```

QY 1 HGHEQOQHGGLGHC--HXKLLDDLEHQGGH 27
    ||| ||| ||| ||| ||| ||| ||| |||
Db 172 HGHVHGCHGCHGSSSHSVSLKQDHGCHGCH 201

RESULT 12
Q93VS6 PRELIMINARY; PRT; 279 AA.
AC Q93VS6;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE At1G67870/T23K23.28 (Unknown protein) (Hypothetical protein).
GN T23K23.28 OR AT1G67870.
OS Arabidopsis thaliana (Mouse-ear cross).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J.,
RA Bowser L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
RA Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,
RA Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,
RA Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,
RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
RA Ecker J.R.;
RT "Arabidopsis cDNA clones.";
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Lam B., Southwick A., Karlin-Neumann G., Nguyen M., Miranda M.,
RA Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,
RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,
RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,
RA Becker J., Theologis A., Davis R.W.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Southwick A., Karlin-Neumann G., Nguyen M., Lam B., Miranda M.,
RA Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,
RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,
RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,
RA Ecker J., Theologis A., Davis R.W.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Kim C.J., Chen H., Cheuk R., Meyers M.C., Shinn P., Banh J.,
RA Bowser L., Carninci P., Chang E., Dale J.M., Goldsmith A.D.,
RA Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Lam B., Lee J.M., Lin J., Miranda M., Narusaka M.,
RA Nguyen M., Onodera C.S., Palm C.J., Quach H.L., Sakurai T., Satou M.,
RA Seki M., Southwick A., Tang C.C., Toriumi M., Wu H.C., Yamada K.,
RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
RA Ecker J.R.;
RT "Arabidopsis ORF clones.";
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RA Haas B.J., Volkovskiy N., Town C.D., Troukhan M., Alexandrov N.,
RA Feldmann K.A., Flavell R.B., White O., Salzberg S.L.;
RT "Full-length messenger RNA sequences greatly improve genome
RT annotation.";
RL Genome Biol. 0:0-0.(2002).
RN [6]
RP SEQUENCE FROM N.A.
RA Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,
RA Feldmann K.;
RT "Full-length cDNA from Arabidopsis thaliana.";
```

```

RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY052272; AAK9465.1; -
DR EMBL; AF386981; AAK62426.1; -
DR EMBL; AY081463; AAM10025.1; -
DR EMBL; AY089864; AAM19974.1; -
DR EMBL; AY085001; AAM61559.1; -
DR GO; GO:0004024; F:alcohol dehydrogenase activity, zinc-dependent; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR InterPro; IPR002328; ADH_ZINC; 1.
DR PROSITE; PS00059; ADH_ZINC; 1.
KW Hypothetical protein_279;
SQ SEQUENCE 279 AA; 30941 MW; 8C9B527458CAC7CF CRC64;

Query Match 40.7%; Score 66; DB 10; Length 279;
Best Local Similarity 46.2%; Pred. No. 0.17;
Matches 12; Conservative 4; Mismatches 10; Indels 0; Gaps 0;

QY 2 GHEQOQGLGCHGXKLLDDLEHQGGH 27
    ||| ||| ||| ||| ||| ||| ||| |||
Db 165 GHGMQGGGCHGCHGCHGCHGCHGCHGCH 190

RESULT 13
Q8CTY5 PRELIMINARY; PRT; 331 AA.
AC Q8CTY5;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Zn-binding lipoprotein adca.
GN SE0193.
OS Staphylococcus epidermidis.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1282;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 12228;
RA Zhang Y., Ren S., Li H., Fu G., Lu L., Lu G., Jia J., Tu Y., Qin Z.,
RA Chen Y., Wen Y.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB016744; AAC03790.1; -
DR GO; GO:0030288; C:periplasmic space (sensu Gram-negative Bact...; IEA.
DR GO; GO:0005488; F:binding; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR006128; Lipoprotein 4.
DR InterPro; IPR000437; Prok_lipoprot_S.
DR InterPro; IPR006127; SBP_bac_9.
DR Pfam; PF01297; SBP_bac_9; 1.
DR PRINTS; PR00690; ADHRSNFAMILY.
DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
KW Lipoprotein; Complete proteome.
SQ SEQUENCE 331 AA; 38007 MW; 1CD02E45A17D200A CRC64;

Query Match 40.7%; Score 66; DB 16; Length 331;
Best Local Similarity 44.0%; Pred. No. 0.2;
Matches 11; Conservative 4; Mismatches 10; Indels 0; Gaps 0;

QY 3 HEQOQGLGCHGXKLLDDLEHQGGH 27
    ||| ||| ||| ||| ||| ||| ||| |||
Db 131 HEHGHGCHGCHGCHGCHGCHGCHGCH 155

RESULT 14
O81050 PRELIMINARY; PRT; 776 AA.
AC O81050;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE At2G03240 protein.
GN Arabidopsis thaliana (Mouse-ear cross).
OS Arabidopsis thaliana; Streptophyta; Embryophyta; Tracheophyta;
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
```


Search completed: March 4, 2004, 13:10:27
Job time : 52.7059 secs

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=20083487; PubMed=10617197;
RA Lin X., Kaut S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblyum T.V.,
Buell C.R., Ketchum K.A., Lee J.J., Roming C.M., Koo H., Moffat K.S.,
Cronin L.A., Shen M., VanAken S.E., Unayam L., Tallon L.J., Gill J.E.,
Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
Copenhaver G.P., Preuss D., Nieman W.C., White O., Eisen J.A.,
Salzberg S.L., Fraser C.M., Venter J.C.;
RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
thaliana.";
RL Nature 402:761-768 (1999).

RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Lin X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AAC005313; AAC34478.1; -.

DR PIR; T02702; T02702.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR InterPro; IPR004342; EXS Cterm.
DR InterPro; IPR004331; SPX.
DR Pfam; PF03124; EXS; A.
DR Pfam; PF03105; SPX; I.
DR PFAM; PF03105; SPX; I.
SQ SEQUENCE 776 AA; 89775 MW; BE30609ACFADD14E CRC64;

Query Match 40.7%; Score 66; DB 10; Length 776;
Best Local Similarity 50.0%; Pred. No. 0.52;
Matches 15; Conservative 1; Mismatches 12; Indels 2; Gaps 1;

Qy 1 HGHEQHQHGLGH--GHXXKLDLDDLEHQGHV 28
Db 76 HGHGCGGGGQGHGPHSDSDDDIEGKGV 105

RESULT 15

Q9NNV9
ID Q9NNV9 PRELIMINARY; PRT; 198 AA.
AC Q9NNV9;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein (fragment).
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=3D7;
RX MEDLINE=21036610; PubMed=11163452;
RA Spielmann T., Beck H.P.;
RT "Analysis of stage-specific transcription in Plasmodium falciparum
RT reveals a set of genes exclusively transcribed in ring stage
RT parasites.";
RL Mol. Biochem. Parasitol. 111:453-458 (2000).
DR EMBL; AJ290925; CAB92934.2; -.
KW Hypothetical protein.
FT NON_TER 1
FT NON_TER 198 198
SQ SEQUENCE 198 AA; 23470 MW; 4DCD4D7E2E4B72 CRC64;

Query Match 40.4%; Score 65.5; DB 5; Length 198;
Best Local Similarity 40.0%; Pred. No. 0.14;
Matches 14; Conservative 1; Mismatches 13; Indels 7; Gaps 1;

Qy 1 HGHEQHQHGLGHGXKXKLDLDDLEH-----QGHHV 28
Db 163 HAEHLDGHGHDGHGHDGHGHDGHGHDGHGHHV 197

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2004, 12:52:17 ; Search time 47.25 Seconds
(without alignments)
101.657 Million cell updates/sec

Title: SEQ-C

Perfect score: 93

Sequence: 1 gkxkLDDLEHOGGHV 17

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Geneseq 29Jan04.*

1: Geneseqp1980s.*

2: Geneseqp1990s.*

3: Geneseqp2000s.*

4: Geneseqp2001s.*

5: Geneseqp2002s.*

6: Geneseqp2003as.*

7: Geneseqp2003bs.*

8: Geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	91	97.8	28	AAV81996	Aay81996 Human hig
2	91	97.8	47	AAV91345	Aay91345 Light cha
3	91	97.8	55	AAV93346	Aay93346 Light cha
4	91	97.8	62	AAV93348	Aay93348 Light cha
5	91	97.8	63	AAV75186	Aar75186 Partial p
6	91	97.8	83	AAV93347	Aay93347 Light cha
7	91	97.8	94	AAV93351	Aay93351 Light cha
8	91	97.8	125	ABE78708	Abb78708 Human hig
9	91	97.8	131	AAV75181	Aar75181 Partial p
10	91	97.8	158	ABE78709	Abb78709 Calmoduli
11	91	97.8	179	AAV93353	Aay93353 Light cha
12	91	97.8	186	AAV93349	Aay93349 Light cha
13	91	97.8	255	AAV93342	Aay93342 Light cha
14	91	97.8	255	AAV73620	Abb73620 Human hig
15	91	97.8	305	ABG21100	Abg21100 Novel hum
16	91	97.8	357	ABR41202	AbR41202 Human DIT
17	91	97.8	415	ABU99146	Abu99146 Novel hum
18	91	97.8	615	ABU99144	Abu99144 Novel hum
19	91	97.8	626	ABU99157	Abu99157 Human hig
20	91	97.8	644	ABG21101	Abg21101 Novel hum
21	91	97.8	644	ABE78710	Abb78710 Human hig
22	91	97.8	644	ABU99150	Abu99150 Novel hum
23	91	97.8	644	ABU99145	Abu99145 Novel hum
24	91	97.8	720	ABG21103	Abg21103 Novel hum
25	87	93.5	16	AAV81998	Aay81998 Human two

ALIGNMENTS

RESULT 1

AAV81996

ID AAV81996 standard; peptide; 28 AA.

XX

AC AAY81996;

XX

DT 16-OCT-2000 (first entry)

XX

DE Human high molecular weight kininogen domain 5 fragment #5.

XX

XX Human; high molecular weight kininogen; HK;

KW two-chain high molecular weight kininogen; HKa; angiogenesis inhibition;

KW tumour; cancer; ocular disorder; rheumatoid arthritis;

KW endothelial cell apoptosis.

XX

OS Homo sapiens.

XX

FN WO200027866-A1.

XX

PD 18-MAY-2000.

XX

PF 05-NOV-1999; 99WO-US026419.

XX

PR 10-NOV-1998; 98US-0107833P.

XX

PA (UTEM) UNIV TEMPLE.

XX

XX (MCCR/) MCCRAE R K.

XX

XX Mcrae RK;

XX

XX WPI; 2000-376483/32.

XX

XX A pharmaceutical composition used to inhibit angiogenesis, inhibit

XX endothelial cell proliferation, and induce endothelial cell apoptosis.

XX

XX Claim 8; Page 28; 52pp; English.

XX

XX The present sequence is derived from human high molecular weight
XX kininogen (HK) domain 5. HK is a 120 kD glycoprotein which binds with
XX high affinity to endothelial cells, where it is cleaved to two-chain high
XX molecular weight kininogen (HKa) by plasma kallikrein. HKa or a synthetic
XX molecular weight kininogen (HKa) by plasma kallikrein. HKa or a synthetic
XX compound comprising the present sequence may be used in a pharmaceutical
XX composition for inhibiting angiogenesis. Angiogenesis occurs in a number
XX of disease states, such as tumour formation and expansion, and certain
XX ocular disorders. It can also occur in a rheumatoid joint, hastening
XX joint destruction by allowing an influx of leukocytes. The composition
XX may inhibit angiogenesis by inhibiting endothelial cell proliferation or
XX by inducing endothelial cell apoptosis. Peptides used in the composition

CC

Fri Mar 5 08:53:31 2004

CC may be recombinant peptides, natural peptides, or synthetic peptides.
CC They may also be chemically synthesised, using, for example, solid phase
CC synthesis methods
XX
SQ Sequence 28 AA;

Query Match 97.8%; Score 91; DB 3; Length 28;
Best Local Similarity 94.1%; Pred. No. 6.4e-08;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GHKKKLDLLEHQGHV 17
||| ||||| ||||| |||||
Db 12 GHKKKLDLLEHQGHV 28

RESULT 2
AAY93345
ID AAY93345 standard; peptide; 47 AA.

XX AC AAY93345;
XX XX 04-SEP-2000 (first entry)
XX XX Light chain of human high molecular weight kininogen fragment.
XX XX Human; high molecular weight kininogen; glycoprotein; endothelial cell;
KW plasma kallikrein; heavy chain; light chain; analogue; angiogenesis;
KW endothelial cell proliferation; endothelial cell migration; vitronectin.
XX XX Synthetic.
OS Homo sapiens.
XX WO200027415-A2.
XX 18-MAY-2000.
XX 09-NOV-1999; 99WO-US026377.
XX 10-NOV-1998; 98US-010784P.
XX (UTEM) UNIV TEMPLE.
PA (DUPO) DUPONT PHARM CO.
PA (COLM/) COLMAN W R.
PA (MOUS/) MOUSA A S.
XX Colman WR, Mousa AS;
XX WPI; 2000-376306/32.
XX Method for inhibiting endothelial cell proliferation, using compound that
PT inhibit endothelial cell migration.
XX Claim 3; Page 36; 41pp; English.

CC The present sequence represents a fragment of the light chain of human
CC high molecular weight kininogen. It is used to produce compounds of the
CC invention. High molecular weight kininogen is a 120 kDa glycoprotein
CC which binds with high affinity to endothelial cells, where it is cleaved
CC by plasma kallikrein into heavy and light chains. Analogues of high
CC molecular weight kininogen are used in the method of the invention. The
CC specification describes a method of inhibiting endothelial cells with a
CC compound containing high molecular weight kininogen analogues. The method
CC and the compounds can be used for inhibiting endothelial cell
CC proliferation. The compounds can also be used for inhibiting
CC angiogenesis. The compounds can also be used to inhibit migration of
CC endothelial cells to vitronectin
XX
SQ Sequence 47 AA;

Query Match 97.8%; Score 91; DB 3; Length 47;
Best Local Similarity 94.1%; Pred. No. 1.1e-07;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GHKKKLDLLEHQGHV 17
||| ||||| ||||| |||||
Db 1 GHKKKLDLLEHQGHV 17

RESULT 3
AAY93346
ID AAY93346 standard; peptide; 55 AA.

XX AC AAY93346;
XX XX 04-SEP-2000 (first entry)
XX XX Light chain of human high molecular weight kininogen analogue.
XX XX Human; high molecular weight kininogen; glycoprotein; endothelial cell;
KW plasma kallikrein; heavy chain; light chain; analogue; angiogenesis;
KW endothelial cell proliferation; endothelial cell migration; vitronectin.
XX XX Synthetic.
OS Homo sapiens.
XX WO200027415-A2.
XX 18-MAY-2000.
XX 09-NOV-1999; 99WO-US026377.
XX 10-NOV-1998; 98US-010784P.
XX (UTEM) UNIV TEMPLE.
PA (DUPO) DUPONT PHARM CO.
PA (COLM/) COLMAN W R.
PA (MOUS/) MOUSA A S.
XX Colman WR, Mousa AS;
XX WPI; 2000-376306/32.
XX Method for inhibiting endothelial cell proliferation, using compound that
PT inhibit endothelial cell migration.
XX Claim 4; Page 36; 41pp; English.

CC The present sequence represents an analogue of the light chain of human
CC high molecular weight kininogen. High molecular weight kininogen is a 120
CC kDa glycoprotein which binds with high affinity to endothelial cells,
CC where it is cleaved by plasma kallikrein into heavy and light chains.
CC Analogues of high molecular weight kininogen are used in the method of
CC the invention. The specification describes a method of inhibiting
CC endothelial cell proliferation. The method comprises contacting
CC endothelial cells with a compound containing high molecular weight
CC kininogen analogues. The method and the compounds can be used for
CC inhibiting endothelial cell proliferation. The compounds can also be used
CC for inhibiting angiogenesis. The compounds can also be used to inhibit
CC migration of endothelial cells to vitronectin
XX
SQ Sequence 55 AA;

Query Match 97.8%; Score 91; DB 3; Length 55;
Best Local Similarity 94.1%; Pred. No. 1.3e-07;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GHKKKLDLLEHQGHV 17
||| ||||| ||||| |||||
Db 37 GHKKKLDLLEHQGHV 53

RESULT 4
AAY93348
ID AAY93348 standard; peptide; 62 AA.

XX AC AAY93348;
XX XX 04-SEP-2000 (first entry)
XX XX Light chain of human high molecular weight kininogen analogue.
XX XX Human; high molecular weight kininogen; glycoprotein; endothelial cell;
KW plasma kallikrein; heavy chain; light chain; analogue; angiogenesis;
KW endothelial cell proliferation; endothelial cell migration; vitronectin.
XX XX Synthetic.
OS Homo sapiens.
XX WO200027415-A2.
XX 18-MAY-2000.
XX 09-NOV-1999; 99WO-US026377.
XX 10-NOV-1998; 98US-010784P.
XX (UTEM) UNIV TEMPLE.
PA (DUPO) DUPONT PHARM CO.
PA (COLM/) COLMAN W R.
PA (MOUS/) MOUSA A S.
XX Colman WR, Mousa AS;
XX WPI; 2000-376306/32.
XX Method for inhibiting endothelial cell proliferation, using compound that
PT inhibit endothelial cell migration.
XX Claim 4; Page 36; 41pp; English.

CC The present sequence represents an analogue of the light chain of human
CC high molecular weight kininogen. High molecular weight kininogen is a 120
CC kDa glycoprotein which binds with high affinity to endothelial cells,
CC where it is cleaved by plasma kallikrein into heavy and light chains.
CC Analogues of high molecular weight kininogen are used in the method of
CC the invention. The specification describes a method of inhibiting
CC endothelial cell proliferation. The method comprises contacting
CC endothelial cells with a compound containing high molecular weight
CC kininogen analogues. The method and the compounds can be used for
CC inhibiting endothelial cell proliferation. The compounds can also be used
CC for inhibiting angiogenesis. The compounds can also be used to inhibit
CC migration of endothelial cells to vitronectin
XX
SQ Sequence 62 AA;

Query Match 97.8%; Score 91; DB 3; Length 62;
Best Local Similarity 94.1%; Pred. No. 1.3e-07;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GHKKKLDLLEHQGHV 17
||| ||||| ||||| |||||
Db 37 GHKKKLDLLEHQGHV 53

RESULT 4
AAY93348
ID AAY93348 standard; peptide; 62 AA.

XX AC AAY93348;
XX XX 04-SEP-2000 (first entry)
XX XX Light chain of human high molecular weight kininogen analogue.
XX XX Human; high molecular weight kininogen; glycoprotein; endothelial cell;
KW plasma kallikrein; heavy chain; light chain; analogue; angiogenesis;
KW endothelial cell proliferation; endothelial cell migration; vitronectin.
XX XX Synthetic.
OS Homo sapiens.
XX WO200027415-A2.
XX 18-MAY-2000.
XX 09-NOV-1999; 99WO-US026377.
XX 10-NOV-1998; 98US-010784P.
XX (UTEM) UNIV TEMPLE.
PA (DUPO) DUPONT PHARM CO.
PA (COLM/) COLMAN W R.
PA (MOUS/) MOUSA A S.
XX Colman WR, Mousa AS;
XX WPI; 2000-376306/32.
XX Method for inhibiting endothelial cell proliferation, using compound that
PT inhibit endothelial cell migration.
XX Claim 4; Page 36; 41pp; English.

CC The present sequence represents an analogue of the light chain of human
CC high molecular weight kininogen. High molecular weight kininogen is a 120
CC kDa glycoprotein which binds with high affinity to endothelial cells,
CC where it is cleaved by plasma kallikrein into heavy and light chains.
CC Analogues of high molecular weight kininogen are used in the method of
CC the invention. The specification describes a method of inhibiting
CC endothelial cell proliferation. The method comprises contacting
CC endothelial cells with a compound containing high molecular weight
CC kininogen analogues. The method and the compounds can be used for
CC inhibiting endothelial cell proliferation. The compounds can also be used
CC for inhibiting angiogenesis. The compounds can also be used to inhibit
CC migration of endothelial cells to vitronectin
XX
SQ Sequence 62 AA;

Query Match 97.8%; Score 91; DB 3; Length 62;
Best Local Similarity 94.1%; Pred. No. 1.3e-07;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GHKKKLDLLEHQGHV 17
||| ||||| ||||| |||||
Db 37 GHKKKLDLLEHQGHV 53

RESULT 4
AAY93348
ID AAY93348 standard; peptide; 62 AA.

AC AAY93348;
 XX
 DT 04-SEP-2000 (first entry)
 XX
 DE Light chain of human high molecular weight kininogen analogue.
 XX
 DE Human; high molecular weight kininogen; glycoprotein; endothelial cell;
 XX plasma kallikrein; heavy chain; light chain; analogue; angiogenesis;
 KW endothelial cell proliferation; endothelial cell migration; vitronectin.
 XX
 XX Synthetic.
 OS Homo sapiens.
 XX
 PN WC200027415-A2.
 XX
 PD 18-MAY-2000.
 XX
 PF 09-NOV-1999; 99WO-US026377.
 XX
 PR 10-NOV-1998; 98US-0107844P.
 XX
 PA (UTEM) UNIV TEMPLE.
 PA (DUPO) DUPONT PHARM CO.
 PA (COLM/) COLMAN W R.
 PA (MOUS/) MOUSA A S.
 XX
 XX Colman WR, Mousa AS;
 PI WPI; 2000-376306/32.
 DR
 XX Method for inhibiting endothelial cell proliferation, using compound that
 PT inhibit endothelial cell migration.
 XX
 PS Claim 6; Page 37; 41pp; English.
 XX
 XX The present sequence represents an analogue of the light chain of human
 CC high molecular weight kininogen. High molecular weight kininogen is a 120
 CC kDa glycoprotein which binds with high affinity to endothelial cells,
 CC where it is cleaved by plasma kallikrein into heavy and light chains.
 CC Analogues of high molecular weight kininogen are used in the method of
 CC the invention. The specification describes a method of inhibiting
 CC endothelial cell proliferation. The method comprises contacting
 CC endothelial cells with a compound containing high molecular weight
 CC kininogen analogues. The method and the compounds can be used for
 CC inhibiting endothelial cell proliferation. The compounds can also be used
 CC for inhibiting angiogenesis. The compounds can also be used to inhibit
 CC migration of endothelial cells to vitronectin
 XX
 XX Sequence 62 AA;
 SQ
 Query Match 97.8%; Score 91; DB 3; Length 62;
 Best Local Similarity 94.1%; Pred. No. 1.5e-07;
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 GHKFKLDDDLHQGHV 17
 DB 16 GHKFKLDDDLHQGHV 32
 RESULT 5
 AAR75186
 ID AAR75186 standard; peptide; 63 AA.
 XX
 AC AAR75186;
 XX
 DT 05-DEC-1995 (first entry)
 XX
 DE Partial peptide of human HMW kininogen fragment 2.
 XX
 DE high molecular weight; kininogen; fragment; 1,2; 1; 2; partial;
 KW wound treating agent; bovine; growth promotion; fibroblast.
 XX
 XX Homo sapiens.
 OS

XX JP07082172-A.
 PN
 XX 28-MAR-1995.
 PD
 PF 17-SEP-1993; 93JP-00230616.
 XX
 PR 17-SEP-1993; 93JP-00230616.
 XX
 PA (FARH) HOECHST JAPAN KK.
 XX
 XX WPI; 1995-158909/21.
 DR
 XX A wound treating agent contg. a partial peptide of kininogen - have
 PT growth promotion activity of fibroblasts.
 XX
 PS Claim 8; Page 8; 8pp; Japanese.
 XX
 CC AAR75186 is a partial peptide corresponding to human kininogen fragment
 CC 1, amino acids 458-520. Partial peptides of bovine and human kininogen
 CC fragments 1,2, 1 and 2, are used in wound treating agent compns. and act
 CC as the active component. The fragments are useful in wound treating
 CC because they have growth promotion activity on fibroblasts
 XX
 XX Sequence 63 AA;
 SQ
 Query Match 97.8%; Score 91; DB 2; Length 63;
 Best Local Similarity 94.1%; Pred. No. 1.5e-07;
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 GHKFKLDDDLHQGHV 17
 DB 17 GHKFKLDDDLHQGHV 33
 RESULT 6
 AAY93347
 ID AAY93347 standard; peptide; 83 AA.
 XX
 AC AAY93347;
 XX
 DT 04-SEP-2000 (first entry)
 XX
 DE Light chain of human high molecular weight kininogen analogue.
 XX
 KW Human; high molecular weight kininogen; glycoprotein; endothelial cell;
 KW plasma kallikrein; heavy chain; light chain; analogue; angiogenesis;
 KW endothelial cell proliferation; endothelial cell migration; vitronectin.
 XX
 XX Synthetic.
 OS Homo sapiens.
 XX
 PN WO200027415-A2.
 XX
 PD 18-MAY-2000.
 XX
 PF 09-NOV-1999; 99WO-US026377.
 XX
 PR 10-NOV-1998; 98US-0107844P.
 XX
 PA (UTEM) UNIV TEMPLE.
 PA (DUPO) DUPONT PHARM CO.
 PA (COLM/) COLMAN W R.
 PA (MOUS/) MOUSA A S.
 XX
 XX Colman WR, Mousa AS;
 PI WPI; 2000-376306/32.
 DR
 XX Method for inhibiting endothelial cell proliferation, using compound that
 PT inhibit endothelial cell migration.
 XX
 PS Claim 5; Page 37; 41pp; English.

Fri Mar 5 08:53:31 2004

inhibiting endothelial cell proliferation. The compounds can also be used for inhibiting angiogenesis. The compounds can also be used to inhibit migration of endothelial cells to vitronectin

Sequence 94 AA;
Query Match 97.8%; Score 91; DB 3; Length 94;
Best Local Similarity 94.1%; Pred. No. 2.1e-07;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GHKXKLDLLEHGGHV 17
| | | | | | | | | | | | | | | | | | | |
DB 37 GHKFKLDDLEHGGHV 53

RESULT 8
ABB78708
ID ABB78708 standard; peptide; 94 AA.
XX AC ABB78708;
XX DT 18-JUL-2002 (first entry)
XX DE Human high molecular weight kininogen (HK) D5 domain SEQ ID NO:2.
XX KW Human; kininogen; high molecular weight kininogen; HK; D5 domain;
KW D5 receptor; angiogenesis; endothelial cell; cytostatic; antitumor;
KW antiatherosclerotic; vasotrophic; vulnerary; tranquilizer; thrombolytic;
KW ophthalmological; gynaecological; antiulcer; antidiabetic; antiarthritic;
KW antiangiogenic; antiapoptotic; endocrine; apoptosis; gene therapy.
XX OS Homo sapiens.
XX PN WO200214369-A2.
XX PD 21-FEB-2002.
XX PF 24-JUL-2001; 2001WO-US023185.
XX PR 24-JUL-2000; 2000US-0220194P.
XX PA (ATTE-) ATTENTON LLC.
XX PI Mazar AP, Juarez JC;
XX DR WPI; 2002-393611/42.
XX DR N-PSDB; ABL52521.
XX PT Novel human kininogen D5 domain polypeptides useful for treating
PT conditions associated with endothelial cell migration, proliferation,
PT invasion or angiogenesis, e.g. arthritis, macular degeneration, benign
PT hyperplasia.
XX PS Claim 2; Page 13; 84pp; English.
XX CC The present invention describes an isolated polypeptide (I) that
CC corresponds to the D5 domain of human kininogen, or biologically active
CC peptide fragment, homologue or functional derivative, and which: (a)
CC inhibits angiogenesis; (b) binds to the D5 binding site on endothelial
CC cells (EC); (c) activates signalling pathways leading to the introduction
CC of apoptosis in EC; and/or (d) inhibits the signalling pathway required
CC for maintenance of EC viability. (I) has cytostatic, antitumor,
CC antiatherosclerotic, vasotrophic, vulnerary, tranquilizer, thrombolytic,
CC ophthalmological, gynaecological, antiulcer, antidiabetic, antiarthritic,
CC antiangiogenic, antiapoptotic and endocrine activities. An antibody (IX)
CC specific for an epitope of (I) is useful for inhibiting tumour growth or
CC angiogenesis in a subject. (I), a D5 fusion polypeptide (II) or a dimeric
CC or trimeric fusion polypeptide (III) can be used for inhibiting EC
CC migration, proliferation, invasion, or angiogenesis, or for inducing EC
CC apoptosis. An angiogenic EC-targeting pharmaceutical composition (X)
CC comprising (I), (II), or (III), can be used for treating a subject having
CC a disease or condition associated with undesired EC migration,
CC proliferation, invasion or angiogenesis. (I), (II), or (III) can be used

The present sequence represents an analogue of the light chain of human high molecular weight kininogen. High molecular weight kininogen is a 120 kDa glycoprotein which binds with high affinity to endothelial cells, where it is cleaved by plasma kallikrein into heavy and light chains. Analogues of high molecular weight kininogen are used in the method of the invention. The specification describes a method of inhibiting endothelial cell proliferation. The method comprises contacting endothelial cells with a compound containing high molecular weight kininogen analogues. The method and the compounds can be used for inhibiting endothelial cell proliferation. The compounds can also be used for inhibiting angiogenesis. The compounds can also be used to inhibit migration of endothelial cells to vitronectin

Sequence 83 AA;
Query Match 97.8%; Score 91; DB 3; Length 83;
Best Local Similarity 94.1%; Pred. No. 2.1e-07;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GHKXKLDLLEHGGHV 17
| | | | | | | | | | | | | | | | | | | |
DB 37 GHKFKLDDLEHGGHV 53

RESULT 7
AAV93351
ID AAV93351 standard; peptide; 94 AA.
XX AC AAV93351;
XX DT 04-SEP-2000 (first entry)
XX DE Light chain of human high molecular weight kininogen analogue.
XX KW Human; high molecular weight kininogen; glycoprotein; endothelial cell;
KW plasma kallikrein; heavy chain; light chain; analogue; angiogenesis;
KW endothelial cell proliferation; endothelial cell migration; vitronectin.
XX OS Synthetic.
XX OS Homo sapiens.
XX PN WO200027415-A2.
XX PD 18-MAY-2000.
XX PF 09-NOV-1999; 99US-US026377.
XX PR 10-NOV-1998; 98US-0107844P.
XX PA (UTEM) UNIV TEMPLE.
PA (DUPO) DUPONT PHARM CO.
PA (COLM/) COLMAN W R.
PA (MOUS/) MOUSA A S.
XX FI Colman WR, Mousa AS;
XX WPI; 2000-376306/32.
XX PT Method for inhibiting endothelial cell proliferation, using compound that
PT inhibit endothelial cell migration.
XX PS Claim 8; Page 39; 41pp; English.
XX CC The present sequence represents an analogue of the light chain of human
CC high molecular weight kininogen. High molecular weight kininogen is a 120
CC kDa glycoprotein which binds with high affinity to endothelial cells,
CC where it is cleaved by plasma kallikrein into heavy and light chains.
CC Analogues of high molecular weight kininogen are used in the method of
CC the invention. The specification describes a method of inhibiting
CC endothelial cell proliferation. The method comprises contacting
CC endothelial cells with a compound containing high molecular weight
CC kininogen analogues. The method and the compounds can be used for

CC for isolating a D5 domain binding molecule from a complex mixture and for
 CC isolating or enriching cells expressing D5 domain binding sites from a
 CC cell mixture. The present sequence represents the human high molecular
 CC weight kininogen (HK) D5 domain amino acid sequence, which is given in
 CC the exemplification of the present invention
 XX
 SQ Sequence 125 AA;

Query Match 97.8%; Score 91; DB 5; Length 125;
 Best Local Similarity 94.1%; Pred. No. 3.3e-07;
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GHKXKLDLLEHOGGHV 17
 ||| ||||| ||||| |||||
 Db 73 GHKFKLDDLEHOGGHV 89

RESULT 9
 AAR75181
 ID AAR75181 standard; peptide; 131 AA.
 XX
 AC AAR75181;
 XX
 DT 05-DEC-1995 (first entry)
 XX
 DE Partial peptide of human HKW kininogen fragment 1.2.
 XX
 KW high molecular weight; kininogen; fragment; 1.2; 1; 2; partial;
 KW wound treating agent; human; growth promotion; fibroblast.
 XX
 OS Homo sapiens.
 XX
 PN JP07082172-A.
 XX
 XX 28-MAR-1995.
 PD
 XX 17-SEP-1993; 93JP-00230616.
 PF
 XX 17-SEP-1993; 93JP-00230616.
 PR
 XX (PARH) HOECHST JAPAN KK.
 PA
 XX WPI; 1995-158909/21.
 DR
 PT A wound treating agent contrg. a partial peptide of kininogen - have
 XX growth promotion activity of fibroblasts.
 XX
 PS Claim 7; Page 7; 8pp; Japanese.
 CC
 CC AAR75181 is a partial peptide corresponding to human kininogen fragment
 CC 1.2, amino acids 350-520. Partial peptides of bovine and human kininogen
 CC fragments 1.2, 1 and 2, are used in wound treating agent compsns. and act
 CC as the active component. The fragments are useful in wound treating
 CC because they have growth promotion activity on fibroblasts
 XX
 SQ Sequence 131 AA;

Query Match 97.8%; Score 91; DB 2; Length 131;
 Best Local Similarity 94.1%; Pred. No. 3.5e-07;
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GHKXKLDLLEHOGGHV 17
 ||| ||||| ||||| |||||
 Db 85 GHKFKLDDLEHOGGHV 101

RESULT 10
 ABB78709
 ID ABB78709 standard; protein; 158 AA.
 XX
 AC ABB78709;
 XX
 DT 18-JUL-2002 (first entry)

Calmodulin binding protein and D5 domain fusion protein SEQ ID NO:13.
 Human; kininogen; high molecular weight kininogen; HK; D5 domain;
 D5 receptor; angiogenesis; endothelial cell; cytotstatic; antitumour;
 antiatherosclerotic; vasotropic; vulnerary; tranquilliser; thrombolytic;
 ophthalmological; gynaecological; antiulcer; antidiabetic; antiarthritic;
 antiangiogenic; antiapoptotic; endocrine; apoptosis; gene therapy;
 calmodulin binding protein; CBP.

Homo sapiens.

Key Location/Qualifiers
 Domain 1..29
 FT /note= "Calmodulin binding protein (CBP) sequence"
 FT 34..158
 FT /label= D5_domain

WO200214369-A2.
 21-FEB-2002.

24-JUL-2001; 2001WO-US023185.

24-JUL-2000; 2000US-0220194P.

(ATTE-) ATTENUON LLC.

Mazar AP, Juarez JC;

WPI; 2002-393611/42.

Novel human kininogen D5 domain polypeptides useful for treating
 conditions associated with endothelial cell migration, proliferation,
 invasion or angiogenesis, e.g. arthritis, macular degeneration, benign
 hyperplasia.

Example 3; Page 62; 84pp; English.

The present invention describes an isolated polypeptide (I) that
 corresponds to the D5 domain of human kininogen, or biologically active
 peptide fragment, homologue or functional derivative, and which: (a)
 inhibits angiogenesis; (b) binds to the D5 binding site on endothelial
 cells (EC); (c) activates signalling pathways leading to the introduction
 of apoptosis in EC; and/or (d) inhibits the signalling pathway required
 for maintenance of EC viability. (I) has cytostatic, antitumour,
 antiatherosclerotic, vasotropic, vulnerary, tranquilliser, thrombolytic,
 ophthalmological, gynaecological, antiulcer, antidiabetic, antiarthritic,
 antiangiogenic, antiapoptotic and endocrine activities. An antibody (IX)
 specific for an epitope of (I) is useful for inhibiting tumour growth or
 angiogenesis in a subject. (II), a D5 fusion polypeptide (II) or a dimeric
 or trimeric fusion polypeptide (III) can be used for inhibiting EC
 migration, proliferation, invasion, or angiogenesis, or for inducing EC
 apoptosis. An angiogenic EC-targeting pharmaceutical composition (X)
 comprising (I), (II), or (III), can be used for treating a subject having
 a disease or condition associated with undesired EC migration,
 proliferation, invasion or angiogenesis. (I), (II), or (III) can be used
 for isolating a D5 domain binding molecule from a complex mixture and for
 isolating or enriching cells expressing D5 domain binding sites from a
 cell mixture. The present sequence represents a calmodulin binding
 protein and D5 domain fusion protein, which is given in an example from
 the present invention

Sequence 158 AA;

Query Match 97.8%; Score 91; DB 5; Length 158;
 Best Local Similarity 94.1%; Pred. No. 4.2e-07;
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GHKXKLDLLEHOGGHV 17
 ||| ||||| ||||| |||||
 Db 106 GHKFKLDDLEHOGGHV 122

```

XX
KW Human; high molecular weight kininogen; glycoprotein; endothelial cell;
KW plasma kallikrein; heavy chain; light chain; analogue; angiogenesis;
KW endothelial cell proliferation; endothelial cell migration; vitronectin.
XX
OS Synthetic.
OS Homo sapiens.
XX
XX WO200027415-A2.
XX
XX 18-MAY-2000.
XX
XX
XX 09-NOV-1999; 99WO-US026377.
XX
XX 10-NOV-1998; 98US-0107844P.
XX
XX (UTEM ) UNIV TEMPLE.
XX (DUPO ) DUPONT PHARM CO.
XX (COLM/) COLMAN W R.
XX (MOUS/) MOUSA A S.
XX
XX Colman WR, Mousa AS;
XX
XX WPI; 2000-376306/32.
XX
XX Method for inhibiting endothelial cell proliferation, using compound that
XX inhibit endothelial cell migration.
XX
XX Claim 9; Page 38; 41pp; English.
XX
XX The present sequence represents an analogue of the light chain of human
XX high molecular weight kininogen. High molecular weight kininogen is a 120
XX kDa glycoprotein which binds with high affinity to endothelial cells,
XX where it is cleaved by plasma kallikrein into heavy and light chains.
XX Analogues of high molecular weight kininogen are used in the method of
XX the invention. The specification describes a method of inhibiting
XX endothelial cell proliferation. The method comprises contacting
XX endothelial cells with a compound containing high molecular weight
XX kininogen analogues. The method and the compounds can be used for
XX inhibiting endothelial cell proliferation. The compounds can also be used
XX for inhibiting angiogenesis. The compounds can also be used to inhibit
XX migration of endothelial cells to vitronectin
XX
XX SQ Sequence 179 AA;
XX
XX Query Match 97.8%; Score 91; DB 3; Length 179;
XX Best Local Similarity 94.1%; Pred. No. 4.9e-07;
XX Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX Qy 1 GHKXKDDLDLEHQGHV 17
XX ||| |||||
XX Db 9 GHKFKLDDLDLEHQGHV 25
XX
XX
XX RESULT 12
XX AAY93349
XX ID AAY93349 standard; peptide; 186 AA.
XX
XX AC AAY93349;
XX
XX
XX 04-SEP-2000 (first entry)
XX
XX DE Light chain of human high molecular weight kininogen analogue.
XX
XX
XX Human; high molecular weight kininogen; glycoprotein; endothelial cell;
KW plasma kallikrein; heavy chain; light chain; analogue; angiogenesis;
KW endothelial cell proliferation; endothelial cell migration; vitronectin.
XX
XX Synthetic.
XX OS Homo sapiens.
XX
XX WO200027415-A2.
XX
XX 18-MAY-2000.
XX
XX
XX 09-NOV-1999; 99WO-US026377.
XX
XX 10-NOV-1998; 98US-0107844P.
XX
XX (UTEM ) UNIV TEMPLE.
XX (DUPO ) DUPONT PHARM CO.
XX (COLM/) COLMAN W R.
XX (MOUS/) MOUSA A S.
XX
XX Colman WR, Mousa AS;
XX
XX WPI; 2000-376306/32.
XX
XX Method for inhibiting endothelial cell proliferation, using compound that
XX inhibit endothelial cell migration.
XX
XX Claim 11; Page 40-41; 41pp; English.
XX
XX The present sequence represents an analogue of the light chain of human
XX high molecular weight kininogen. High molecular weight kininogen is a 120
XX kDa glycoprotein which binds with high affinity to endothelial cells,
XX where it is cleaved by plasma kallikrein into heavy and light chains.
XX Analogues of high molecular weight kininogen are used in the method of
XX the invention. The specification describes a method of inhibiting
XX endothelial cell proliferation. The method comprises contacting
XX endothelial cells with a compound containing high molecular weight
XX kininogen analogues. The method and the compounds can be used for
XX inhibiting endothelial cell proliferation. The compounds can also be used
XX for inhibiting angiogenesis. The compounds can also be used to inhibit
XX migration of endothelial cells to vitronectin
XX
XX SQ Sequence 179 AA;
XX
XX Query Match 97.8%; Score 91; DB 3; Length 179;
XX Best Local Similarity 94.1%; Pred. No. 4.9e-07;
XX Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX Qy 1 GHKXKDDLDLEHQGHV 17
XX ||| |||||
XX Db 9 GHKFKLDDLDLEHQGHV 25
XX
XX
XX RESULT 12
XX AAY93349
XX ID AAY93349 standard; peptide; 186 AA.
XX
XX AC AAY93349;
XX
XX
XX 04-SEP-2000 (first entry)
XX
XX DE Light chain of human high molecular weight kininogen analogue.

```

```

XX
KW Human; high molecular weight kininogen; glycoprotein; endothelial cell;
KW plasma kallikrein; heavy chain; light chain; analogue; angiogenesis;
KW endothelial cell proliferation; endothelial cell migration; vitronectin.
XX
OS Synthetic.
OS Homo sapiens.
XX
XX WO200027415-A2.
XX
XX 18-MAY-2000.
XX
XX
XX 09-NOV-1999; 99WO-US026377.
XX
XX 10-NOV-1998; 98US-0107844P.
XX
XX (UTEM ) UNIV TEMPLE.
XX (DUPO ) DUPONT PHARM CO.
XX (COLM/) COLMAN W R.
XX (MOUS/) MOUSA A S.
XX
XX Colman WR, Mousa AS;
XX
XX WPI; 2000-376306/32.
XX
XX Method for inhibiting endothelial cell proliferation, using compound that
XX inhibit endothelial cell migration.
XX
XX Claim 9; Page 38; 41pp; English.
XX
XX The present sequence represents an analogue of the light chain of human
XX high molecular weight kininogen. High molecular weight kininogen is a 120
XX kDa glycoprotein which binds with high affinity to endothelial cells,
XX where it is cleaved by plasma kallikrein into heavy and light chains.
XX Analogues of high molecular weight kininogen are used in the method of
XX the invention. The specification describes a method of inhibiting
XX endothelial cell proliferation. The method comprises contacting
XX endothelial cells with a compound containing high molecular weight
XX kininogen analogues. The method and the compounds can be used for
XX inhibiting endothelial cell proliferation. The compounds can also be used
XX for inhibiting angiogenesis. The compounds can also be used to inhibit
XX migration of endothelial cells to vitronectin
XX
XX SQ Sequence 186 AA;
XX
XX Query Match 97.8%; Score 91; DB 3; Length 186;
XX Best Local Similarity 94.1%; Pred. No. 5.1e-07;
XX Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX Qy 1 GHKXKDDLDLEHQGHV 17
XX ||| |||||
XX Db 16 GHKFKLDDLDLEHQGHV 32
XX
XX
XX RESULT 13
XX AAY93342
XX ID AAY93342 standard; protein; 255 AA.
XX
XX AC AAY93342;
XX
XX
XX 04-SEP-2000 (first entry)
XX
XX DE Light chain of human high molecular weight kininogen.
XX
XX
XX Human; high molecular weight kininogen; glycoprotein; endothelial cell;
KW plasma kallikrein; heavy chain; light chain; analogue; angiogenesis;
KW endothelial cell proliferation; endothelial cell migration; vitronectin.
XX
XX Synthetic.
XX OS Homo sapiens.
XX
XX WO200027415-A2.
XX
XX 18-MAY-2000.

```

XX 09-NOV-1999; 99WC-US026377.
 PF 10-NOV-1998; 98US-0107844P.
 PR (UTEM) UNIV TEMPLE.
 PA (DUPO) DUPONT PHARM CO.
 PA (COLM/) COLMAN W R.
 PA (MOUS/) MOUSA A S.
 XX Colman WR, Mousa AS;
 FT WPI; 2000-376306/32.
 DR Method for inhibiting endothelial cell proliferation, using compound that
 FT inhibit endothelial cell migration.
 XX Disclosure; Page 3; 41pp; English.
 PS The present sequence represents the light chain of human high molecular
 CC weight kininogen. High molecular weight kininogen is a 120 kDa
 CC glycoprotein which binds with high affinity to endothelial cells, where
 CC it is cleaved by plasma kallikrein into heavy and light chains. Analogues
 CC of high molecular weight kininogen are used in the method of the
 CC invention. The specification describes a method of inhibiting endothelial
 CC cell proliferation. The method comprises contacting endothelial cells
 CC with a compound containing high molecular weight kininogen analogues. The
 CC method and the compounds can be used for inhibiting endothelial cell
 CC proliferation. The compounds can also be used for inhibiting
 CC angiogenesis. The compounds can also be used to inhibit migration of
 CC endothelial cells to vitronectin
 XX Sequence 255 AA;
 SQ

Query Match 97.8%; Score 91; DB 3; Length 255;
 Best Local Similarity 94.1%; Pred. No. 7.2e-07;
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GHKXKLDLDDLEHOGGHV 17
 DB 85 GHKFKLDLDDLEHOGGHV 101

RESULT 14
 AAB73620
 ID AAB73620 standard; protein; 255 AA.
 XX AAB73620;
 AC AAB73620;
 XX 10-AUG-2001 (first entry)
 DT Human high molecular weight kininogen (HK) light chain.
 DE Human; high molecular weight kininogen; HK; light chain; domain 5;
 KW antibody; monoclonal antibody C1C1; neovascularisation inhibition;
 KW endothelial cell proliferation inhibition; antiangiogenic;
 KW vascular tube formation inhibition; diabetic retinopathy;
 KW rheumatoid arthritis; atherosclerotic plaque rupture; cancer; tumour;
 KW cystostatic; antidiabetic; ophthalmological; antirheumatic; antiarthritic;
 KW antiatherosclerotic.
 XX Homo sapiens.
 OS

XX Key Location/Qualifiers
 FH 1..12
 FT Region /note= "C-terminal portion of HK domain 4 remaining after
 FT bradykinin liberation"
 FT 13..131
 FT Domain /label= Domain 5
 FT /note= "Corresponds to residues 384-502 of HK"
 FT 69..131
 FT Region /note= "Referred to in Claim 3; corresponds to HK
 FT residues 440-502"

FT Region 69..84
 FT /note= "Referred to in Claim 4; corresponds to HK
 FT residues 440-457"
 FT 70..86
 FT Domain /label= His/Gly-rich_subdomain
 FT 100..125
 FT Domain /label= Endothelial_cell_binding_domain
 FT 104..131
 FT Domain /label= His/Gly/lys-rich_subdomain
 FT 115..131
 FT Region /note= "Referred to in Claim 5; corresponds to HK
 FT residues 486-502"
 FT 132..255
 FT Domain /label= Domain 6
 FT /note= "Corresponds to residues 503-626 of HK"
 XX WO200134195-A1.
 PN 17-MAY-2001.
 XX 10-NOV-2000; 2000WO-US030975.
 XX 12-NOV-1999; 99US-0165165P.
 PR (UTEM) UNIV TEMPLE.
 PA (DUPO) DUPONT PHARM CO.
 XX Colman RW, Mousa SA;
 PI WPI; 2001-328940/34.
 DR Inhibiting angiogenesis in a mammal using an antibody against high
 FT molecular weight kininogen domain 5.
 XX Claim 4; Page 2; 38pp; English.
 PS The invention relates to a method of inhibiting angiogenesis in a mammal,
 CC comprising administering an antibody against an epitope of high molecular
 CC weight kininogen (HK) domain 5. In particular, the antibody used can be
 CC monoclonal antibody C1C1 which is produced by hybridoma ATCC HB-8964.
 CC The method of the invention is used to inhibit endothelial cell
 CC proliferation, vascular tube formation and/or neovascularisation in
 CC disease states such as diabetic retinopathy, rheumatoid arthritis and
 CC atherosclerotic plaques. The antibody may be administered to prevent
 CC plaque rupture, which leads to thrombotic occlusion of coronary or
 CC cerebral arteries. The antibody may also be used to inhibit tumour growth
 CC via the inhibition of angiogenesis. The present sequence represents human
 CC high molecular weight kininogen (HK) light chain, which contains HK
 CC domain 5
 XX Sequence 255 AA;
 SQ

Query Match 97.8%; Score 91; DB 4; Length 255;
 Best Local Similarity 94.1%; Pred. No. 7.2e-07;
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GHKXKLDLDDLEHOGGHV 17
 DB 85 GHKFKLDLDDLEHOGGHV 101

RESULT 15
 ABG21100
 ID ABG21100 standard; protein; 305 AA.
 XX ABG21100;
 AC ABG21100;
 XX 18-FEB-2002 (first entry)
 DT Novel human diagnostic protein #21091.
 DE Human; chromosome mapping; gene mapping; gene therapy; forensic;
 XX food supplement; medical imaging; diagnostic; genetic disorder.
 KW

Fri Mar 5 08:53:31 2004

```

XX OS Homo sapiens.
XX PN WO200175067-A2.
XX PD 11-OCT-2001.
XX PF 30-MAR-2001; 2001WO-US008631.
XX PR 31-MAR-2000; 2000US-00540217.
XX PR 23-AUG-2000; 2000US-00649167.
XX PA (HYSE-) HYSEQ INC.
XX PI Drmanac RT, Liu C, Tang YT;
XX DR N-PSDB; AAS85287.
XX PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX FS Claim 20; SEQ ID NO 51459; 103pp; English.
XX CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 305 AA;

```

```

Query Match 97.8%; Score 91; DB 4; Length 305;
Best Local Similarity 94.1%; Pred. No. 8.7e-07;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 GHKKLDDDLLEHGGHV 17
DB 135 GHKKLDDDLLEHGGHV 151

```

Search completed: March 4, 2004, 13:07:27
Job time : 48.25 secs

Result No.	Score	Query Match	Length	DB	ID	Description	
1	91	97.8	47	3	US-09-612-126-4	Sequence 4, Appli	
2	91	97.8	55	3	US-09-612-126-5	Sequence 5, Appli	
3	91	97.8	62	3	US-09-612-126-7	Sequence 7, Appli	
4	91	97.8	83	3	US-09-612-126-6	Sequence 6, Appli	
5	91	97.8	94	3	US-09-612-126-10	Sequence 10, Appl	
6	91	97.8	179	3	US-09-612-126-11	Sequence 11, Appl	
7	91	97.8	186	3	US-09-612-126-8	Sequence 8, Appli	
8	91	97.8	255	3	US-09-612-126-1	Sequence 1, Appli	
9	46	49.5	270	4	US-09-252-991A-20626	Sequence 20626, A	
10	44	47.3	97	4	US-09-621-976-6625	Sequence 6625, Ap	
11	44	47.3	106	2	US-08-820-925-14	Sequence 14, Appl	
12	44	47.3	106	3	US-09-307-917-14	Sequence 14, Appl	
13	44	47.3	106	4	US-09-734-036-14	Sequence 14, Appl	
14	44	47.3	107	1	US-08-409-731A-2	Sequence 2, Appli	
15	44	47.3	107	2	US-08-470-298B-2	Sequence 2, Appli	
16	44	47.3	107	2	US-09-023-073A-2	Sequence 2, Appli	
17	44	47.3	107	3	US-09-361-737-2	Sequence 2, Appli	
18	44	47.3	135	2	US-08-820-925-2	Sequence 2, Appli	
19	44	47.3	135	3	US-08-899-031-1	Sequence 1, Appli	
20	44	47.3	135	3	US-09-307-917-2	Sequence 2, Appli	
21	44	47.3	135	4	US-09-734-036-2	Sequence 2, Appli	
22	44	47.3	139	4	US-09-621-976-4574	Sequence 4574, Ap	
23	44	47.3	725	4	US-09-252-991A-24201	Sequence 24201, A	
24	43	46.2	524	4	US-09-540-236-2706	Sequence 2706, Ap	
25	43	46.2	765	4	US-09-252-991A-17416	Sequence 17416, A	
26	42	45.2	378	4	US-09-252-991A-22754	Sequence 22754, A	
27	42	45.2	762	4	US-09-252-991A-28078	Sequence 28078, A	

seq-C.rai

Fri Mar 5 08:53:32 2004

```

; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 55
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Human high
; OTHER INFORMATION: molecular weight kininogen light chain amino acids
; OTHER INFORMATION: Lys(420) through Asp(474)
US-09-612-126-5

Query Match          97.8%; Score 91; DB 3; Length 55;
Best Local Similarity 94.1%; Pred. No. 3.8e-08;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GHKXKLDLLEHGGHV 17
   ||| ||||| ||||| |||||
Db 37 GHKFKLDDLEHGGHV 53

RESULT 3
US-09-612-126-7
; Sequence 7, Application US/09612126
; Patent No. 6284726
; GENERAL INFORMATION:
; APPLICANT: TEMPLE UNIVERSITY - OF THE COMMONWEALTH SYSTEM OF
; TITLE OF INVENTION: INHIBITION OF ANGIOGENESIS BY PEPTIDE ANALOGS OF HIGH
; TITLE OF INVENTION: MOLECULAR WEIGHT KININOGEN DOMAN 5
; FILE REFERENCE: 6056-258 CT1
; CURRENT APPLICATION NUMBER: US/09/612,126
; CURRENT FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/107,844
; PRIOR FILING DATE: 1998-11-10
; PRIOR APPLICATION NUMBER: PCT/US99/26377
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 62
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Human high
; OTHER INFORMATION: molecular weight kininogen light chain amino acids
; OTHER INFORMATION: His(441) through Lys(502)
US-09-612-126-7

Query Match          97.8%; Score 91; DB 3; Length 62;
Best Local Similarity 94.1%; Pred. No. 4.3e-08;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GHKXKLDLLEHGGHV 17
   ||| ||||| ||||| |||||
Db 16 GHKFKLDDLEHGGHV 32

RESULT 4
US-09-612-126-6
; Sequence 6, Application US/09612126
; Patent No. 6284726
; GENERAL INFORMATION:
; APPLICANT: TEMPLE UNIVERSITY - OF THE COMMONWEALTH SYSTEM OF
; TITLE OF INVENTION: INHIBITION OF ANGIOGENESIS BY PEPTIDE ANALOGS OF HIGH
; TITLE OF INVENTION: MOLECULAR WEIGHT KININOGEN DOMAN 5
; FILE REFERENCE: 6056-258 CT1
; CURRENT APPLICATION NUMBER: US/09/612,126
; CURRENT FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/107,844
; PRIOR FILING DATE: 1998-11-10
; PRIOR APPLICATION NUMBER: PCT/US99/26377
; PRIOR FILING DATE: 1999-11-09

```

```

; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 83
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Human high
; OTHER INFORMATION: molecular weight kininogen light chain amino acids
; OTHER INFORMATION: Lys(420) through Lys(502)
US-09-612-126-6

Query Match          97.8%; Score 91; DB 3; Length 83;
Best Local Similarity 94.1%; Pred. No. 6e-08;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GHKXKLDLLEHGGHV 17
   ||| ||||| ||||| |||||
Db 37 GHKFKLDDLEHGGHV 53

RESULT 5
US-09-612-126-10
; Sequence 10, Application US/09612126
; Patent No. 6284726
; GENERAL INFORMATION:
; APPLICANT: TEMPLE UNIVERSITY - OF THE COMMONWEALTH SYSTEM OF
; TITLE OF INVENTION: INHIBITION OF ANGIOGENESIS BY PEPTIDE ANALOGS OF HIGH
; TITLE OF INVENTION: MOLECULAR WEIGHT KININOGEN DOMAN 5
; FILE REFERENCE: 6056-258 CT1
; CURRENT APPLICATION NUMBER: US/09/612,126
; CURRENT FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/107,844
; PRIOR FILING DATE: 1998-11-10
; PRIOR APPLICATION NUMBER: PCT/US99/26377
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 94
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Human high
; OTHER INFORMATION: molecular weight kininogen light chain amino acids
; OTHER INFORMATION: Lys(420) through Ser(513)
US-09-612-126-10

Query Match          97.8%; Score 91; DB 3; Length 94;
Best Local Similarity 94.1%; Pred. No. 6.9e-08;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GHKXKLDLLEHGGHV 17
   ||| ||||| ||||| |||||
Db 37 GHKFKLDDLEHGGHV 53

RESULT 6
US-09-612-126-11
; Sequence 11, Application US/09612126
; Patent No. 6284726
; GENERAL INFORMATION:
; APPLICANT: TEMPLE UNIVERSITY - OF THE COMMONWEALTH SYSTEM OF
; TITLE OF INVENTION: INHIBITION OF ANGIOGENESIS BY PEPTIDE ANALOGS OF HIGH
; TITLE OF INVENTION: MOLECULAR WEIGHT KININOGEN DOMAN 5
; FILE REFERENCE: 6056-258 CT1
; CURRENT APPLICATION NUMBER: US/09/612,126
; CURRENT FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/107,844
; PRIOR FILING DATE: 1998-11-10
; PRIOR APPLICATION NUMBER: PCT/US99/26377
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 12

```

```

; SEQ ID NO 1
; LENGTH: 255
; TYPE: PRT
; ORGANISM: Human
US-09-612-126-1

Query Match          97.8%; Score 91; DB 3; Length 255;
Best Local Similarity 94.1%; Pred.No. 2.1e-07;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 GHKXKLDLLEHQGHV 17
DB      85 GHKFLDLDLEHQGHV 101
      ||| ||||| ||||| |||
      ||| ||||| ||||| |||

RESULT 9
US-09-252-991A-20626
; Sequence 20626, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 20626
; LENGTH: 270
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-20626

Query Match          49.5%; Score 46; DB 4; Length 270;
Best Local Similarity 50.0%; Pred.No. 5.2;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY      2 HKXKLDLLEHQGHV 17
DB      154 HQHQLDQDQHQGHV 169
      ||| ||||| ||||| |||
      ||| ||||| ||||| |||

RESULT 10
US-09-621-976-6625
; Sequence 6625, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 6625
; LENGTH: 97
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-621-976-6625

Query Match          47.3%; Score 44; DB 4; Length 97;
Best Local Similarity 53.8%; Pred.No. 3.6;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY      5 KLDDLLEHQGHV 17
DB      38 KPDKEIHQGNEM 50
      ||| ||||| ||||| |||
      ||| ||||| ||||| |||

```

— 6 —

; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 106 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-734-036-14

Query Match 47.3%; Score 44; DB 4; Length 106;
Best Local Similarity 53.8%; Pred. No. 4;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 5 KLDDDLHQGHV 17
DB 37 KPDKEIEHQGNHM 49

RESULT 14

US-08-409-731A-2
; Sequence 2, Application US/08409731A

; Patent No. 5658758

; GENERAL INFORMATION:

; APPLICANT: Ni, Jian

; APPLICANT: Yu, Guo-Liang

; APPLICANT: Gentz, Reiner

; APPLICANT: Rosen, Craig A.

; TITLE OF INVENTION: CYTOSTATIN I

; NUMBER OF SEQUENCES: 11

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: HUMAN GENOME SCIENCES, INC.

; STREET: 9410 KEY WEST AVENUE

; CITY: ROCKVILLE

; STATE: MD

; COUNTRY: USA

; ZIP: 20850

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/409,731A

; FILING DATE: 24-MAR-1995

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Benson, Robert H

; REGISTRATION NUMBER: 30,446

; REFERENCE/DOCKET NUMBER: PF175

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 301-309-8504

; TELEFAX: 301-309-8512

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 107 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-409-731A-2

Query Match 47.3%; Score 44; DB 1; Length 107;
Best Local Similarity 53.8%; Pred. No. 4;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 5 KLDDDLHQGHV 17
DB 38 KPDKEIEHQGNHM 50

RESULT 15

US-08-470-298B-2

; Sequence 2, Application US/08470298B

; Patent No. 5844061

; GENERAL INFORMATION:
; APPLICANT: NI, JIAN
; APPLICANT: GENTZ, REINER
; APPLICANT: YU, GUO-LIANG
; APPLICANT: ROSEN, CRAIG A.
; TITLE OF INVENTION: CYTOSTATIN I
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HUMAN GENOME SCIENCES, INC.
; STREET: 9410 KEY WEST AVENUE
; CITY: ROCKVILLE
; STATE: MD
; COUNTRY: US
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/470,298B
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BROOKES, ALLAN A.
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PF175D1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 301-309-8504
; TELEFAX: 301-309-8512
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 107 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-470-298B-2

Query Match 47.3%; Score 44; DB 2; Length 107;
Best Local Similarity 53.8%; Pred. No. 4;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 5 KLDDDLHQGHV 17
DB 38 KPDKEIEHQGNHM 50

Search completed: March 4, 2004, 13:12:26
Job time : 13 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2004, 13:10:38 ; Search time 25.75 Seconds
(without alignments)
139.402 Million cell updates/sec

Title: SEQ-C

Perfect score: 93

Sequence: 1 ghxxkLDDLEHGGGHV 17

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 809742 seqs, 211153259 residues

Total number of hits satisfying chosen parameters: 809742

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:

- 1: /cgn2_6/ptodata/2/pubaa/PCT_NEW_PUB.pep.*
- 2: /cgn2_6/ptodata/2/pubaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubaa/US07_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/2/pubaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	91	97.8	415	15	US-10-162-335-76
2	91	97.8	615	15	US-10-162-335-72
3	91	97.8	644	15	US-10-162-335-74
4	91	97.8	644	15	US-10-162-335-84
5	44	47.3	70	9	US-09-737-149-46
6	44	47.3	70	9	US-09-737-149-47
7	44	47.3	106	13	US-10-153-740-14
8	44	47.3	107	9	US-09-901-436A-2
9	44	47.3	135	9	US-09-737-149-16
10	44	47.3	135	9	US-09-737-149-48
11	44	47.3	135	9	US-09-737-149-49
12	44	47.3	135	13	US-10-153-740-2
13	44	47.3	135	13	US-10-032-366-1
14	44	47.3	156	9	US-09-737-149-14
15	44	47.3	344	14	US-10-156-761-11768

Sequence 9568, Ap
Sequence 6977, Ap
Sequence 72, Appl
Sequence 1930, Ap
Sequence 43, Appl
Sequence 5703, Ap
Sequence 5834, Ap
Sequence 1296, A
Sequence 1261, A
Sequence 3556, Ap
Sequence 8556, Ap
Sequence 1606, Ap
Sequence 20518, A
Sequence 20, Appl
Sequence 2138, Ap
Sequence 3141, Ap
Sequence 7, Appl
Sequence 8141, Ap
Sequence 373, App
Sequence 1325, Ap
Sequence 74, Appl
Sequence 72, Appl
Sequence 1, Appl
Sequence 110, App
Sequence 3, Appl
Sequence 28, Appl
Sequence 7707, Ap
Sequence 19846, A
Sequence 98, Appl
Sequence 99, Appl

ALIGNMENTS

RESULT 1

US-10-162-335-76

Sequence 76, Application US/10162335

Publication No. US20040009480A1

GENERAL INFORMATION:

APPLICANT: Anderson, David W.
APPLICANT: Baumgartner, Jason C.
APPLICANT: Boldog, Ferenc L.
APPLICANT: Casman, Stacie J.
APPLICANT: Edinger, Shlomit R.
APPLICANT: Gangolli, Esha A.
APPLICANT: Gerlach, Valerie
APPLICANT: Gorman, Linda
APPLICANT: Guo, Xiaojia (Sasha)
APPLICANT: Hjalte, Ford
APPLICANT: Kekuda, Ramesh
APPLICANT: Li, Li
APPLICANT: MacDougall, John R.
APPLICANT: Malyankar, Uriel M.
APPLICANT: Miller, Isabelle
APPLICANT: Padigaru, Muralidhara
APPLICANT: Patturajan, Meera
APPLICANT: Pena, Carol E. A.
APPLICANT: Rastelli, Luca
APPLICANT: Shimkets, Richard A.
APPLICANT: Stone, David J.
APPLICANT: Spytek, Kimberly A.
APPLICANT: Vernet, Corine A. M.
APPLICANT: Voss, Edward Z.
APPLICANT: Zerhusen, Bryan D.
TITLE OF INVENTION: Therapeutic Polypeptides, Nucleic Acids Encoding Same, and Method
FILE REFERENCE: 21402-377 B
CURRENT APPLICATION NUMBER: US/10/162,335
CURRENT FILING DATE: 2002-10-01
PRIOR APPLICATION NUMBER: 60/295,607
PRIOR FILING DATE: 2001-06-04
PRIOR APPLICATION NUMBER: 60/295,661
PRIOR FILING DATE: 2001-06-04

; PRIOR APPLICATION NUMBER: 60/296,404
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: 60/296,418
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: 60/297,414
; PRIOR FILING DATE: 2001-06-11
; PRIOR APPLICATION NUMBER: 60/297,567
; PRIOR FILING DATE: 2001-06-12
; PRIOR APPLICATION NUMBER: 60/298,285
; PRIOR FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: 60/298,556
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/299,949
; PRIOR FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: 60/300,883
; PRIOR FILING DATE: 2001-06-26
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 201
; SEQ ID NO 76
; LENGTH: 415
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-162-335-72

Query Match 97.8%; Score 91; DB 15; Length 415;
Best Local Similarity 94.1%; Pred. No. 4.1e-06;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GHXXKLDLLEHGGHV 17
||| ||||| ||||| |||||
DB 245 GHKFKLDDLEHGGHV 261

RESULT 2
US-10-162-335-72
; Sequence 72, Application US/10162335
; Publication No. US20040009480A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, David W.
; APPLICANT: Baumgartner, Jason C.
; APPLICANT: Boldog, Ferenc L.
; APPLICANT: Casman, Stacie J.
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Gangolli, Esha A.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Gorman, Linda
; APPLICANT: Hjal, Tord
; APPLICANT: Guo, Xiaojia (Sasha)
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Li, Li
; APPLICANT: MacDougall, John R.
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Millet, Isabelle
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Patturajan, Meera
; APPLICANT: Pena, Carol E. A.
; APPLICANT: Rastelli, Luca
; APPLICANT: Shinkets, Richard A.
; APPLICANT: Stone, David J.
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Vernet, Corine A. M.
; APPLICANT: Voss, Edward Z.
; APPLICANT: Zerhusen, Bryan D.
; TITLE OF INVENTION: Therapeutic Polypeptides, Nucleic Acids Encoding Same, and Method
; FILE REFERENCE: 21402-377 B
; CURRENT APPLICATION NUMBER: US/10/162,335
; CURRENT FILING DATE: 2002-10-01
; PRIOR APPLICATION NUMBER: 60/295,607
; PRIOR FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: 60/295,661
; PRIOR FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: 60/296,404
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: 60/296,418
; PRIOR FILING DATE: 2001-06-06

; PRIOR APPLICATION NUMBER: 60/296,418
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: 60/297,414
; PRIOR FILING DATE: 2001-06-11
; PRIOR APPLICATION NUMBER: 60/297,567
; PRIOR FILING DATE: 2001-06-12
; PRIOR APPLICATION NUMBER: 60/298,285
; PRIOR FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: 60/298,556
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/299,949
; PRIOR FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: 60/300,883
; PRIOR FILING DATE: 2001-06-26
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 201
; SEQ ID NO 72
; LENGTH: 615
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-162-335-72

Query Match 97.8%; Score 91; DB 15; Length 615;
Best Local Similarity 94.1%; Pred. No. 6.3e-06;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GHXXKLDLLEHGGHV 17
||| ||||| ||||| |||||
DB 445 GHKFKLDDLEHGGHV 461

RESULT 3
US-10-162-335-74
; Sequence 74, Application US/10162335
; Publication No. US20040009480A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, David W.
; APPLICANT: Baumgartner, Jason C.
; APPLICANT: Boldog, Ferenc L.
; APPLICANT: Casman, Stacie J.
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Gangolli, Esha A.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Gorman, Linda
; APPLICANT: Hjal, Tord
; APPLICANT: Guo, Xiaojia (Sasha)
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Li, Li
; APPLICANT: MacDougall, John R.
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Millet, Isabelle
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Patturajan, Meera
; APPLICANT: Pena, Carol E. A.
; APPLICANT: Rastelli, Luca
; APPLICANT: Shinkets, Richard A.
; APPLICANT: Stone, David J.
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Vernet, Corine A. M.
; APPLICANT: Voss, Edward Z.
; APPLICANT: Zerhusen, Bryan D.
; TITLE OF INVENTION: Therapeutic Polypeptides, Nucleic Acids Encoding Same, and Methods
; FILE REFERENCE: 21402-377 B
; CURRENT APPLICATION NUMBER: US/10/162,335
; CURRENT FILING DATE: 2002-10-01
; PRIOR APPLICATION NUMBER: 60/295,607
; PRIOR FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: 60/295,661
; PRIOR FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: 60/296,404
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: 60/296,418
; PRIOR FILING DATE: 2001-06-06


```

; PRIOR APPLICATION NUMBER: 60/297,414
; PRIOR FILING DATE: 2001-06-11
; PRIOR APPLICATION NUMBER: 60/297,567
; PRIOR FILING DATE: 2001-06-12
; PRIOR APPLICATION NUMBER: 60/298,285
; PRIOR FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: 60/298,556
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/299,949
; PRIOR FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: 60/300,883
; PRIOR FILING DATE: 2001-06-26
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 201
; SEQ ID NO 74
; LENGTH: 644
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-162-335-74

Query Match          97.8%; Score 91; DB 15; Length 644;
Best Local Similarity 94.1%; Pred. No. 6.6e-06;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GHXXKLDLDEHQGHV 17
   ||| ||||| ||||| |||||
Db 474 GHXFKLDDDLDEHQGHV 490

RESULT 4
US-10-162-335-84
; Sequence 84, Application US/10162335
; Publication No. US20040009480A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, David W.
; APPLICANT: Baumgartner, Jason C.
; APPLICANT: Boldog, Ferenc L.
; APPLICANT: Casman, Stacie J.
; APPLICANT: Eninger, Shalomit R.
; APPLICANT: Gangolli, Esha A.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Gorman, Linda
; APPLICANT: Guo, Xiaojia (Sasha)
; APPLICANT: Hjal, Tord
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Li, Li
; APPLICANT: MacDougall, John R.
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Millet, Isabelle
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Patturajan, Meera
; APPLICANT: Pena, Carol E. A.
; APPLICANT: Restelli, Luca
; APPLICANT: Shimkets, Richard A.
; APPLICANT: Stone, David J.
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Vernet, Corine A. M.
; APPLICANT: Voss, Edward Z.
; APPLICANT: Zerhusen, Bryan D.
; TITLE OF INVENTION: Therapeutic Polypeptides, Nucleic Acids Encoding Same, and Method
; FILE REFERENCE: 21402-377 B
; CURRENT APPLICATION NUMBER: US/10162,335
; CURRENT FILING DATE: 2002-10-01
; PRIOR APPLICATION NUMBER: 60/295,607
; PRIOR FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: 60/295,661
; PRIOR FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: 60/296,404
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: 60/296,418
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: 60/297,414
; PRIOR FILING DATE: 2001-06-11

```

```

; PRIOR APPLICATION NUMBER: 60/297,567
; PRIOR FILING DATE: 2001-06-12
; PRIOR APPLICATION NUMBER: 60/298,285
; PRIOR FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: 60/298,556
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/299,949
; PRIOR FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: 60/300,883
; PRIOR FILING DATE: 2001-06-26
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 201
; SEQ ID NO 84
; LENGTH: 644
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-162-335-84

Query Match          97.8%; Score 91; DB 15; Length 644;
Best Local Similarity 94.1%; Pred. No. 6.6e-06;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GHXXKLDLDEHQGHV 17
   ||| ||||| ||||| |||||
Db 474 GHXFKLDDDLDEHQGHV 490

RESULT 5
US-09-737-149-46
; Sequence 46, Application US/09737149
; Patent No. US20020077466A1
; GENERAL INFORMATION:
; APPLICANT: Spaderna, Steven K
; APPLICANT: Quinn, Kerry E.
; APPLICANT: Shimkets, Richard A.
; APPLICANT: Muralidhara, Padigaru
; APPLICANT: Spytek, Kimberly A.
; TITLE OF INVENTION: Polypeptides and Nucleic Acids Encoding Same
; FILE REFERENCE: 15966-620 CIP
; CURRENT APPLICATION NUMBER: US/09737,149
; CURRENT FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/170,564
; PRIOR FILING DATE: 1999-12-14
; PRIOR APPLICATION NUMBER: 60/173,165
; PRIOR FILING DATE: 1999-12-27
; PRIOR APPLICATION NUMBER: 60/173,362
; PRIOR FILING DATE: 1999-12-27
; PRIOR APPLICATION NUMBER: 60/173,544
; PRIOR FILING DATE: 1999-12-29
; PRIOR APPLICATION NUMBER: 60/174,404
; PRIOR FILING DATE: 2000-01-04
; PRIOR APPLICATION NUMBER: 60/174,962
; PRIOR FILING DATE: 2000-01-07
; PRIOR APPLICATION NUMBER: 60/223,929
; PRIOR FILING DATE: 2000-08-09
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 46
; LENGTH: 70
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-737-149-46

Query Match          47.3%; Score 44; DB 9; Length 70;
Best Local Similarity 53.8%; Pred. No. 12;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 5 KLDDDLDEHQGHV 17
   ||| ||||| |||||
Db 38 KPDKEIEHQGNM 50

RESULT 6

```

US-09-737-149-47
; Sequence 47, Application US/09737149
; Patent No. US20020077466A1
; GENERAL INFORMATION:
; APPLICANT: Spaderna, Steven X
; APPLICANT: Quinn, Kerry E.
; APPLICANT: Shimkets, Richard A.
; APPLICANT: Muralidhara, Padigaru
; APPLICANT: Szytek, Kimberly A.
; TITLE OF INVENTION: Polypeptides and Nucleic Acids Encoding Same
; FILE REFERENCE: 15966-620 CIP
; CURRENT APPLICATION NUMBER: US/09/737,149
; CURRENT FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/170,564
; PRIOR FILING DATE: 1999-12-14
; PRIOR APPLICATION NUMBER: 60/173,165
; PRIOR FILING DATE: 1999-12-27
; PRIOR APPLICATION NUMBER: 60/173,362
; PRIOR FILING DATE: 1999-12-27
; PRIOR APPLICATION NUMBER: 60/173,544
; PRIOR FILING DATE: 1999-12-29
; PRIOR APPLICATION NUMBER: 60/174,404
; PRIOR FILING DATE: 2000-01-04
; PRIOR APPLICATION NUMBER: 60/174,962
; PRIOR FILING DATE: 2000-01-07
; PRIOR APPLICATION NUMBER: 60/223,929
; PRIOR FILING DATE: 2000-08-09
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 47
; LENGTH: 70
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)..(70)
; OTHER INFORMATION: Where X is a residue at which the query and
; other information: subject sequences are not identical.
US-09-737-149-47

Query Match 47.3%; Score 44; DB 9; Length 70;
Best Local Similarity 53.8%; Pred. No. 12;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 5 KLDDLEHQGHV 17
DB 38 KPDKEHQGNM 50

RESULT 7
US-10-153-740-14
; Sequence 14, Application US/10153740
; Publication No. US20020147149A1
; GENERAL INFORMATION:
; APPLICANT: NI, JIAN
; YU, GUO-LIANG
; GENTZ, REINER L.
; DILLON, PATRICK
; TITLE OF INVENTION: CYTOSTATIN III
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HUMAN GENOME SCIENCES, INC.
; STREET: 9410 KEY WEST AVENUE
; CITY: ROCKVILLE
; STATE: MD
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/153,740
; FILING DATE: 24-May-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/734,036
; FILING DATE: <Unknown>
; APPLICATION NUMBER: 09/307,817
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: BROOKES, ANDERS A.
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PF222
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 106 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 14:
US-10-153-740-14

Query Match 47.3%; Score 44; DB 13; Length 106;
Best Local Similarity 53.8%; Pred. No. 19;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 5 KLDDLEHQGHV 17
DB 37 KPDKEHQGNM 49

RESULT 8
US-09-901-436A-2
; Sequence 2, Application US/09901436A
; Patent No. US20020098515A1
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: Cytostatin I
; FILE REFERENCE: PFI75D4
; CURRENT APPLICATION NUMBER: US/09/901,436A
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 09/361,737
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: 09/023,073
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: 08/470,298
; PRIOR FILING DATE: 1995-06-06
; PRIOR APPLICATION NUMBER: 08/409,731
; PRIOR FILING DATE: 1995-03-24
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 107
; TYPE: PRT
; ORGANISM: human
US-09-901-436A-2

Query Match 47.3%; Score 44; DB 9; Length 107;
Best Local Similarity 53.8%; Pred. No. 19;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 5 KLDDLEHQGHV 17
DB 38 KPDKEHQGNM 50

RESULT 9
US-09-737-149-16
; Sequence 16, Application US/09737149

```
; Patent No. US20020077466A1
; GENERAL INFORMATION:
; APPLICANT: Spaderna, Steven K
; APPLICANT: Quinn, Kerry E.
; APPLICANT: Shimkets, Richard A.
; APPLICANT: Muralidhara, Padigaru
; APPLICANT: Spytek, Kimberly A.
; TITLE OF INVENTION: Polypeptides and Nucleic Acids Encoding Same
; FILE REFERENCE: 15966-620 CIP
; CURRENT APPLICATION NUMBER: US/09/737,149
; CURRENT FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/170,564
; PRIOR FILING DATE: 1999-12-14
; PRIOR APPLICATION NUMBER: 60/173,165
; PRIOR FILING DATE: 1999-12-27
; PRIOR APPLICATION NUMBER: 60/173,362
; PRIOR FILING DATE: 1999-12-27
; PRIOR APPLICATION NUMBER: 60/174,404
; PRIOR FILING DATE: 1999-12-29
; PRIOR APPLICATION NUMBER: 60/174,962
; PRIOR FILING DATE: 2000-01-04
; PRIOR APPLICATION NUMBER: 60/223,929
; PRIOR FILING DATE: 2000-08-09
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16
; LENGTH: 135
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-737-149-16

Query Match          47.3%; Score 44; DB 9; Length 135;
Best Local Similarity 53.8%; Pred. No. 24;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY      5 KLDDDLHQGHV 17
        |||:|||||:
Db      38 KPDKEIHHQGNM 50

RESULT 10
US-09-737-149-48
; Sequence 48, Application US/09737149
; Patent No. US20020077466A1
; GENERAL INFORMATION:
; APPLICANT: Spaderna, Steven K
; APPLICANT: Quinn, Kerry E.
; APPLICANT: Shimkets, Richard A.
; APPLICANT: Muralidhara, Padigaru
; APPLICANT: Spytek, Kimberly A.
; TITLE OF INVENTION: Polypeptides and Nucleic Acids Encoding Same
; FILE REFERENCE: 15966-620 CIP
; CURRENT APPLICATION NUMBER: US/09/737,149
; CURRENT FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/170,564
; PRIOR FILING DATE: 1999-12-14
; PRIOR APPLICATION NUMBER: 60/173,165
; PRIOR FILING DATE: 1999-12-27
; PRIOR APPLICATION NUMBER: 60/173,362
; PRIOR FILING DATE: 1999-12-27
; PRIOR APPLICATION NUMBER: 60/174,404
; PRIOR FILING DATE: 2000-01-04
; PRIOR APPLICATION NUMBER: 60/174,962
; PRIOR FILING DATE: 2000-01-07
; PRIOR APPLICATION NUMBER: 60/223,929
; PRIOR FILING DATE: 2000-08-09
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16
; LENGTH: 135
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-737-149-16

Query Match          47.3%; Score 44; DB 9; Length 135;
Best Local Similarity 53.8%; Pred. No. 24;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY      5 KLDDDLHQGHV 17
        |||:|||||:
Db      38 KPDKEIHHQGNM 50

RESULT 10
US-09-737-149-48
; Sequence 48, Application US/09737149
; Patent No. US20020077466A1
; GENERAL INFORMATION:
; APPLICANT: Spaderna, Steven K
; APPLICANT: Quinn, Kerry E.
; APPLICANT: Shimkets, Richard A.
; APPLICANT: Muralidhara, Padigaru
; APPLICANT: Spytek, Kimberly A.
; TITLE OF INVENTION: Polypeptides and Nucleic Acids Encoding Same
; FILE REFERENCE: 15966-620 CIP
; CURRENT APPLICATION NUMBER: US/09/737,149
; CURRENT FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/170,564
; PRIOR FILING DATE: 1999-12-14
; PRIOR APPLICATION NUMBER: 60/173,165
; PRIOR FILING DATE: 1999-12-27
; PRIOR APPLICATION NUMBER: 60/173,362
; PRIOR FILING DATE: 1999-12-27
; PRIOR APPLICATION NUMBER: 60/174,404
; PRIOR FILING DATE: 2000-01-04
; PRIOR APPLICATION NUMBER: 60/174,962
; PRIOR FILING DATE: 2000-01-07
; PRIOR APPLICATION NUMBER: 60/223,929
; PRIOR FILING DATE: 2000-08-09
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16
```

```
; LENGTH: 135
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-737-149-48

Query Match          47.3%; Score 44; DB 9; Length 135;
Best Local Similarity 53.8%; Pred. No. 24;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY      5 KLDDDLHQGHV 17
        |||:|||||:
Db      38 KPDKEIHHQGNM 50

RESULT 11
US-09-737-149-49
; Sequence 49, Application US/09737149
; Patent No. US20020077466A1
; GENERAL INFORMATION:
; APPLICANT: Spaderna, Steven K
; APPLICANT: Quinn, Kerry E.
; APPLICANT: Shimkets, Richard A.
; APPLICANT: Muralidhara, Padigaru
; APPLICANT: Spytek, Kimberly A.
; TITLE OF INVENTION: Polypeptides and Nucleic Acids Encoding Same
; FILE REFERENCE: 15966-620 CIP
; CURRENT APPLICATION NUMBER: US/09/737,149
; CURRENT FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/170,564
; PRIOR FILING DATE: 1999-12-14
; PRIOR APPLICATION NUMBER: 60/173,165
; PRIOR FILING DATE: 1999-12-27
; PRIOR APPLICATION NUMBER: 60/173,362
; PRIOR FILING DATE: 1999-12-27
; PRIOR APPLICATION NUMBER: 60/173,544
; PRIOR FILING DATE: 1999-12-29
; PRIOR APPLICATION NUMBER: 60/174,404
; PRIOR FILING DATE: 2000-01-04
; PRIOR APPLICATION NUMBER: 60/174,962
; PRIOR FILING DATE: 2000-01-07
; PRIOR APPLICATION NUMBER: 60/223,929
; PRIOR FILING DATE: 2000-08-09
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 49
; LENGTH: 135
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Consensus Sequence
; NAME/KEY: VARIANT
; LOCATION: (1)..(133)
; OTHER INFORMATION: Where X represents a residue at which the query
; OTHER INFORMATION: and subject sequences are not identical.
US-09-737-149-49

Query Match          47.3%; Score 44; DB 9; Length 135;
Best Local Similarity 53.8%; Pred. No. 24;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY      5 KLDDDLHQGHV 17
        |||:|||||:
Db      38 KPDKEIHHQGNM 50

RESULT 12
US-10-153-740-2
; Sequence 2, Application US/10153740
; Publication No. US20020147149A1
; GENERAL INFORMATION:
; APPLICANT: NI, JIAN
; APPLICANT: YU, GUO-LIANG
; APPLICANT: GENZY, REINER L.
```


QY 5 KLDDLEHGGHV 17
 Db 38 KPDKEHGGNHM 50

RESULT 15

US-10-156-761-11768
 ; Sequence 11768, Application US/10156761
 ; Publication No. US20030119018A1
 ; GENERAL INFORMATION:
 ; APPLICANT: OMURA, SATOSHI
 ; APPLICANT: IKEDA, HARUO
 ; APPLICANT: ISHIKAWA, JUN
 ; APPLICANT: HORIKAWA, HIROSHI
 ; APPLICANT: SHIBA, TADAYOSHI
 ; APPLICANT: SAKAKI, YOSHIYUKI
 ; APPLICANT: HATTORI, MASAHIRA
 ; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
 ; FILE REFERENCE: 249-262
 ; CURRENT APPLICATION NUMBER: US/10/156,761
 ; CURRENT FILING DATE: 2002-05-29
 ; PRIOR APPLICATION NUMBER: JP 2001-204089
 ; PRIOR FILING DATE: 2001-05-30
 ; PRIOR APPLICATION NUMBER: JP 2001-272697
 ; PRIOR FILING DATE: 2001-08-02
 ; NUMBER OF SEQ ID NOS: 15109
 ; SEQ ID NO 11768
 ; LENGTH: 344
 ; TYPE: PRT
 ; ORGANISM: Streptomyces avermitilis
 US-10-156-761-11768

Query Match 47.3%; Score 44; DB 14; Length 344;
 Best Local Similarity 53.3%; Pred. No. 65;
 Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 GHKXKLDLLEHGG 15
 Db 254 GSYTKLDDDMVYDGG 268

Search completed: March 4, 2004, 13:26:33
 Job time : 27 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2004, 13:02:33 ; Search time 10.5 seconds
(without alignments)
155.739 Million cell updates/sec

Title: SEQ-C
Perfect score: 93
Sequence: 1 ghkxkLDDLEHQGHV 17

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues
Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_78:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	91	97.8	644	1 KGHUHI	kininogen, HMW pre
2	47	50.5	541	2 P34850	probable acid-CoA
3	46	49.5	487	2 E30251	aspartate oxidase
4	43	46.2	174	2 G37112	hypothetical prote
5	43	46.2	489	2 T36100	probable ATP-bind
6	43	46.2	683	2 S01433	repressor protein
7	43	46.2	1878	2 E86189	hypothetical prote
8	42.5	45.7	290	2 C27115	K-kininogen, HMW p
9	42.5	45.7	315	2 A27115	major acute phase
10	42.5	45.7	536	2 T27668	hypothetical prote
11	42	45.2	125	2 G38017	hypothetical prote
12	42	45.2	210	2 G97120	uncharacterized co
13	42	45.2	237	2 T27198	hypothetical prote
14	42	45.2	245	2 T23844	hypothetical prote
15	42	45.2	621	1 K3E0H1	kininogen, HMW I p
16	42	45.2	849	2 E61631	probable membrane
17	42	45.2	1061	2 S37667	trac-1 protein - E
18	42	45.2	1448	2 T08526	DNA primase trac2
19	42	45.2	1448	2 S37669	trac-2 protein - E
20	42	45.2	1785	2 S53976	probable membrane
21	42	45.2	3282	2 E82750	hemagglutinin-like
22	42	45.2	3442	2 E82589	hemagglutinin-like
23	42	45.2	3455	2 E82519	K-kininogen, HMW p
24	41.5	44.6	264	2 C25486	kininogen, HMW I p
25	41.5	44.6	639	2 A25486	cytochrome P450 52
26	41	44.1	538	1 O4CK34	genome polyprotein
27	41	44.1	2332	1 GNNY4F	hypothetical prote
28	41	44.1	6713	2 E89921	hypothetical prote
29	40.5	43.5	924	2 S34926	hypothetical prote

ALIGNMENTS

RESULT 1

KGHUHI
Kininogen, HMW precursor [validated] - human
N:Alternate names: alpha-2-thiol proteinase inhibitor; preprokininogen; prokininogen
N:Contains: bradykinin (kallidin I); HMW kininogen I; HMW kininogen II; low molecular wei
C:Species: Homo sapiens (man)
C>Date: 28-May-1986 #sequence revision 28-May-1986 #text change 08-Dec-2000
C:Accession: A01279; A25276; S32422; A91153; A24871; A27899; A31905; A34030; S024
R:Ohkubo, I.; Kurachi, K.; Takasawa, T.; Shiohara, H.; Sasaki, M.
Biochemistry 23, 5691-5697, 1984
A>Title: Isolation of a human cDNA for alpha-2-thiol proteinase inhibitor and its identit
A:Reference number: A90490; MUID:85122621; PMID:6441591
A:Accession: A01273
A:Molecule type: mRNA
A:Residues: 1-389 <OHK>
A:Cross-references: GB:K02566; NID:g177889
R:Takagaki, Y.; Kitamura, N.; Nakanishi, S.
J. Biol. Chem. 260, 8601-8609, 1985
A>Title: Cloning and sequence analysis of cDNAs for human high molecular weight and low m
A:Reference number: A92544; MUID:85234582; PMID:2989293
A:Accession: A25276
A:Molecule type: mRNA
A:Residues: 1-592, 'I', 594-644 <TAK>
A:Cross-references: GB:M1437; NID:g186751; PIDN:AAB59550.1; PID:G386852
R:Auerwald, E.A.; Roessler, D.; Mentele, R.; Assfalg-Machleidt, I.
FEBS Lett. 321, 93-97, 1993
A>Title: Cloning, expression and characterization of human kininogen domain 3.
A:Reference number: S32422; MUID:86030270; PMID:8467916
A:Accession: S32422
A:Molecule type: mRNA
A:Residues: 'ANSM', 253-377 <AUE>
A>Note: differences are due to known cloning artifacts
R:Lottspeich, F.; Kellermann, J.; Henschen, A.; Foertsch, B.; Muller-Esterl, W.
Eur. J. Biochem. 152, 307-314, 1985
A>Title: The amino acid sequence of the light chain of human high-molecular-mass kininoge
A:Reference number: A91153; MUID:86030270; PMID:4054110
A:Accession: A91153
A:Molecule type: protein
A:Residues: 379-644 <LOI>
A>Note: the bradykinin sequence preceding the light chain sequence was not determined in
R:Lottspeich, F.; Kellermann, J.; Lottspeich, F.; Henschen, A.; Mueller-Esterl, W.
Eur. J. Biochem. 154, 471-478, 1986
A>Title: Completion of the primary structure of human high-molecular-mass kininogen. The
A:Reference number: A24871; MUID:86108361; PMID:3484703
A:Accession: A24871
A:Molecule type: protein
A:Residues: 'Z', 20-380 <KEI1>
R:Lottspeich, F.; Kellermann, J.; Lottspeich, F.; Henschen, A.; Mueller-Esterl, W.
in Kinins IV, Greenbaum, L.M., ed., pp.85-89, Plenum Press, New York
A>Title: Amino acid sequence of the light chain of human high molecular mass kininogen.
A:Reference number: A27899
A:Accession: A27899

hypothetical prote
zinc finger protei
citrate (pro-38)-1
protein L homolog
protein T17H7.13 {
hypothetical prote
puff C-8 protein -
conserved hypotet
valine-tRNA ligase
fibronectin-bindin
sodium bicarbonate
sodium bicarbonate co
Na+ bicarbonate co
sodium bicarbonate
tensin - chicken
tensin, cardiac mu

30 40 43.0 91 2 C81931
31 40 43.0 239 2 I38615
32 40 43.0 356 2 P82277
33 40 43.0 414 2 H4231
34 40 43.0 415 2 B6434
35 40 43.0 453 2 T17823
36 40 43.0 457 2 S39079
37 40 43.0 725 2 B82425
38 40 43.0 878 2 F64425
39 40 43.0 1025 2 S69790
40 40 43.0 1035 2 T13962
41 40 43.0 1035 2 T14110
42 40 43.0 1079 2 PC7034
43 40 43.0 1079 2 T14031
44 40 43.0 1733 2 S27939
45 40 43.0 1744 2 A54970

A:Molecule type: protein
A:Residues: 379-389, K, 390-407, Q, 409-644 <KSL2>
R:Minidrou, T.; Carretero, O.A.; Proust, D.; Walz, D.; Scicchi, A.G.
Biochem. Biophys. Res. Commun. 152, 519-526, 1988
A:Title: A new kinin molecule in human plasma kininogens.
A:Reference number: A27699; PMID:88209021; PMID:3365237
A:Accession: A27699
A:Molecule type: protein
A:Residues: 380-389 <WIK>
R:Maeda, H.; Matsumura, Y.; Kato, H.
J. Biol. Chem. 263, 16051-16054, 1988
A:Title: Purification and identification of [hydroxyprolyl(3)]bradykinin in ascitic fluid
A:Reference number: A31905; PMID:89034061; PMID:3182782
A:Accession: A31905
A:Molecule type: protein
A:Residues: 381-389 <MAE>
R:Sasaguri, M.; Ikeda, M.; Ideishi, M.; Arakawa, K.
Biochem. Biophys. Res. Commun. 150, 511-516, 1988
A:Title: Identification of [hydroxyproline(3)]-lysyl-bradykinin released from human plasma
A:Reference number: A34030; PMID:88106632; PMID:3337729
A:Accession: A34030
A:Molecule type: protein
A:Residues: 380-389 <SAS>
R:Lenarcic, B.; Gabrijelcic, D.; Rozman, B.; Drobnic-Kosorok, M.; Turk, V.
Biol. Chem. Hoppe-Seyler 369, 257-261, 1988
A:Title: Human cathepsin B and cysteine proteinase inhibitors (CPis) in inflammatory and
A:Reference number: S02482; PMID:89076517; PMID:3264507
A:Accession: S02482
A:Molecule type: protein
A:Residues: 1-19;189-192;310-314;381-389 <LENI>
R:Kato, H.; Matsumura, Y.; Maeda, H.
FEBS Lett. 232, 252-254, 1988
A:Title: Isolation and identification of hydroxyproline analogues of bradykinin in human
A:Reference number: A61495; PMID:88211869; PMID:3366244
A:Accession: A61495
A:Molecule type: protein
A:Residues: 380-389 <KAT1>
A:Experimental source: urine
A:Note: this peptide had Pro-383 modified to 4-hydroxyproline
A:Accession: S61495
A:Molecule type: protein
A:Residues: 381-389 <KAT2>
A:Experimental source: urine
A:Note: this peptide had Pro-383 modified to 4-hydroxyproline
A:Accession: C61495
A:Molecule type: protein
A:Residues: 380-389 <KAT3>
R:Lenarcic, B.; Krasovec, M.; Ritonja, A.; Olafsson, I.; Turk, V.
FEBS Lett. 280, 211-215, 1991
A:Title: Inactivation of human cystatin C and kininogen by human cathepsin D.
A:Reference number: S14303; PMID:91192133; PMID:2013314
A:Accession: S14447
A:Molecule type: protein
A:Residues: 264-359, N, 361-375 <LEN2>
R:Little, S.S.; Johnson, D.A.
Biochem. J. 307, 341-346, 1995
A:Title: Human mast cell tryptase isoforms: separation and examination of substrate-specific
A:Reference number: S55239; PMID:95251593; PMID:7733867
A:Accession: S55239
A:Molecule type: protein
A:Residues: 450-452, X, 454, X, 456 <LIT>
R:Straczek, J.; Maachi, F.; le Nguyen, D.; Becchi, M.; Heulin, M.H.; Nabet, P.; Bellevil
FEBS Lett. 373, 207-211, 1995
A:Title: Purification from human plasma of a tetrapeptide that potentiates insulin-like
A:Reference number: S68059; PMID:96033974; PMID:7589467
A:Accession: S68059
A:Molecule type: protein
A:Residues: 431-434 <STR>
R:Kitamura, N.; Kitagawa, H.; Fukushima, D.; Takagaki, Y.; Miyata, T.; Nakanishi, S.
J. Biol. Chem. 260, 8610-8617, 1985
A:Title: Structural organization of the human kininogen gene and a model for its evolution
A:Reference number: A92545; PMID:85234583; PMID:2989294
A:Contents: annotation; gene organization

R:Pierce, J.V.
Fed. Proc. 27, 52-57, 1968
A:Title: Structural features of plasma kinins and kininogens.
A:Reference number: A91455; PMID:90255622; PMID:4952632
A:Contents: annotation; bradykinin
C:Comment: The HMW kininogen precursor and the LMW form are produced from the same gene
C:Comment: Kininogen is a cysteine proteinase inhibitor, takes part in initiation of the
C:Comment: The glycine/histidine/lysine-rich region of HMW kininogen light chain is impo
C:Comment: Bradykinin, released from kininogen by kallikrein, is a potent vasodilator, ir
xyproline residue is present in the kininogen prior to the release of bradykinin.
C:Genetics:
A:Gene: GDB:XNG
A:Cross-references: GDB:125256; OMIM:228960
A:Map position: 3q27-3q27
A:Introns: 65/3; 102/3; 131/1; 188/3; 224/3; 253/1; 310/3; 346/3; 375/3
C:Superfamily: kininogen; cystatin homology
C:Keywords: alternative splicing; blood coagulation; cysteine proteinase inhibitor; dupli
F:1-18/Domain: signal sequence #status experimental <Sig>
F:19-644/Product: HMW kininogen I (prokininogen) #status experimental <WAT1>
F:19-379,390-644/Product: HMW kininogen II #status experimental <WAT2>
F:19-379/Domain: HMW kininogen heavy chain #status experimental <HCH>
F:19-131/Domain: cystatin homology <CY1>
F:142-253/Domain: cystatin homology <CY2>
F:264-375/Domain: cystatin homology <CY3>
F:380-389/Product: lysyl-bradykinin (kallidin II) #status experimental <KBDY>
F:381-389/Product: bradykinin (kallidin I) #status experimental <BDY>
F:390-644/Domain: HMW kininogen light chain #status experimental <LCH>
F:421-510/Region: glycine/histidine/lysine-rich 30-residue repeats
F:431-434/Product: low molecular weight growth promoting factor #status experimental <GPI
F:19/Modified site: pyrrolidone carboxylic acid (Glu) (in mature form) #status experimental
F:28-614,83-94,107-126,142-145,206-218,229-248,264-267,328-340,351-370/Disulfide bonds: #
F:169,205,294/Binding site: carboxylate (Asn) (covalent) #status experimental
F:379-380/Cleavage site: Met-Lys (kallikrein) #status experimental
F:383/Modified site: 4-hydroxyproline (Pro) (partial) #status experimental
F:389-390/Cleavage site: Arg-Ser (kallikrein) #status experimental
F:401,533,542,546,557,571,593,628/Binding site: carboxylate (Thr) (covalent) #status ex
F:577/Binding site: carboxylate (Ser) (covalent) #status experimental

Query Match 97.8%; Score 91; DB 1; Length 644;
Best Local Similarity 94.1%; Pred. No. 7.7e-07;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GHXXKLDLLEHQGHV 17
||| |||||
Db 474 GHKFKLDLLEHQGHV 490

RESULT 2

T34850
Probable acid-CoA ligase (EC 6.2.1.-) SC2G5.17 [similarity] - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C:Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 18-Aug-2000
C:Accession: T34850
C:Oliver, K.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, February 1999
A:Reference number: Z21559
A:Accession: T34850
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-541 <OLI>
A:Cross-references: EMBL:AL035478; PIDN:CAB36604.1; GSPDB:GN00070; SCOEDB:SC2G5.17
A:Experimental source: strain A3(2)
C:Genetics:
A:Gene: SCOEDB:SC2G5.17
C:Superfamily: 4-coumarate-CoA ligase; acetate-CoA ligase homology <ACL>
C:Keywords: acid-thiol ligase
F:70-533/Domain: acetate-CoA ligase homology <ACL>

Query Match 50.5%; Score 47; DB 2; Length 541;
Best Local Similarity 53.3%; Pred. No. 9.1;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 3 KXKLDLLEHQGHV 17
: : : : :
Db 346 QTRMDLLEHRTGV 360

RESULT 3
E90251
aspartate oxidase [nadb] [imported] - Sulfolobus solfataricus
C:Species: Sulfolobus solfataricus
C:Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 15-Jun-2001
C:Accession: E90251
R:Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, R.
arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.
submitted to GenBank, April 2001
A:Description: Sulfolobus solfataricus complete genome.
A:Reference number: A99139
A:Accession: E90251
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-487 <KUR>
A:Cross-references: GB:AE006641; NID:g13814182; PIDN:AAK41268.1; GSPDB:GN00155
C:Genetics:
A:Gene: nadB
C:Superfamily: fumarate reductase flavoprotein; 3-oxosteroid 1-dehydrogenase homology; 5

Query Match 49.5%; Score 46; DB 2; Length 487;
Best Local Similarity 66.7%; Pred. No. 12;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 5 KLDDLEHQGHV 16
: : : : :
Db 109 KFDDDLLEGGH 120

RESULT 4
G83712
hypothetical protein BH0503 [imported] - Bacillus halodurans (strain C-125)
C:Species: Bacillus halodurans
C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C:Accession: G83712
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A:Reference number: A83650; MUID:20512582; PMID:11058132
A:Accession: G83712
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-174 <STO>
A:Cross-references: GB:AP001508; GB:BA000004; NID:g10172890; PIDN:BA04222.1; GSPDB:GN00
A:Experimental source: strain C-125
C:Genetics:
A:Gene: BH0503
C:Superfamily: Deinococcus radiodurans hypothetical protein DR0763

Query Match 46.2%; Score 43; DB 2; Length 174;
Best Local Similarity 58.3%; Pred. No. 12;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 6 LDDLEHQGHV 17
: : : : :
Db 88 LNDMLHERGGHI 99

RESULT 5
T36100
probable ATP-binding protein - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C:Accession: T36100
R:Saunders, D.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, April 1999
A:Reference number: Z11596

A:Accession: T36100
A:Status: preliminary; translated from GE/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-489 <SAU>
A:Cross-references: EMBL:AL049661; PIDN:CAB41216.1; GSPDB:GN00070; SCODEB:SCE134.18
A:Experimental source: strain A3(2)
C:Genetics:
A:Gene: SCODEB:SCE134.18

Query Match 46.2%; Score 43; DB 2; Length 489;
Best Local Similarity 47.1%; Pred. No. 36;
Matches 8; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

Qy 1 GHKXKLDLLEHQGHV 17
: : : : :
Db 262 GSPERLDQALGHQAGHL 278

RESULT 6
S01433
repressor protein C - phage phi-C31
C:Species: phage phi-C31
C:Date: 30-Sep-1989 #sequence_revision 30-Sep-1989 #text_change 04-Mar-2000
C:Accession: S01433; S38912
R:Sinclair, R.B.; Bibb, M.J.
Mol. Gen. Genet. 213, 269-277, 1988
A:Title: The repressor gene (c) of the Streptomyces temperate phage phi-c31: nucleotide s
A:Reference number: S01433; MUID:89039715; PMID:3185504
A:Accession: S01433
A:Molecule type: DNA
A:Residues: 1-683 <SIN>
A:Cross-references: EMBL:X12865; NID:g15458; PIDN:CAA31345.1; PID:g15459
R:Hartley, N.M.; Murphy, G.O.; Bruton, C.J.; Chater, K.F.
submitted to the EMBL Data Library, November 1993
A:Reference number: S38912
A:Accession: S38912
A:Molecule type: DNA
A:Residues: 1-683 <HAR>
A:Cross-references: EMBL:X76288; NID:g432610; PIDN:CAA53911.1; PID:g432611
C:Genetics:
A:Gene: c
C:Superfamily: phage phi-C31 repressor protein C
C:Keywords: DNA binding; transcription regulation

Query Match 46.2%; Score 43; DB 2; Length 683;
Best Local Similarity 70.0%; Pred. No. 52;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 7 DDLLEHQGHV 16
: : : : :
Db 474 DDDVERQGAH 483

RESULT 7
E86189
hypothetical protein [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cross)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: E86189
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chen, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Li, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, I
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: E86189
A:Status: preliminary
A:Molecule type: DNA

A;Residues: 1-1878 <STO>
 A;Cross-references: GB:AE005172; NID:94836907; PIDN:AAD30609.1; GSPDB:GN00141
 C;Genetics:
 A;Map position: 1

Query Match 46.2%; Score 43; DB 2; Length 1878;
 Best Local Similarity 70.0%; Pred. No. 1.5e+02;
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 GHXXKLDDDL 10
 |||||
 Db 224 GHKKLDEDI 233

RESULT 8

C27115
 K-kininogen, LMW precursor - rat (fragments)
 C;Species: Rattus norvegicus (Norway rat)
 C;Date: 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change 20-Aug-1999
 C;Accession: C27115, A25488
 R;Fung, W.P.; Schreiber, G.
 J. Biol. Chem. 262, 9298-9308, 1987
 A;Title: Structure and expression of the genes for major acute phase alpha-1-protein (th
 A;Reference number: A92653; MUID:87250580; PMID:2439509
 A;Accession: C27115
 A;Molecule type: DNA
 A;Residues: 1-230 <FUN>
 R;Kagayama, R.; Kitamura, N.; Ohkubo, H.; Nakanishi, S.
 J. Biol. Chem. 262, 2345-2351, 1987
 A;Title: Differing utilization of homologous transcription initiation sites of rat K and
 A;Reference number: A25488; MUID:87137465; PMID:3818598
 A;Accession: A25488
 A;Status: Preliminary
 A;Molecule type: DNA
 A;Residues: 1-48 <KAG>
 A;Cross-references: GB:J02662; NID:g205071; PIDN:AAA41483.1; PID:g205072
 C;Superfamily: kininogen, cystatin homology
 F;19-65/Domain: cystatin homology (fragment) <CVS>

Query Match 45.7%; Score 42.5; DB 2; Length 290;
 Best Local Similarity 40.0%; Pred. No. 25;
 Matches 10; Conservative 1; Mismatches 5; Indels 9; Gaps 1;

QY 1 GHXXKLDD-----DLHQGGH 16
 |||||
 Db 135 GHKKLDDLKQORDGYNRHPMGH 159

RESULT 9

A27115
 major acute phase alpha-1 protein 1 - rat (fragments)
 C;Species: Rattus norvegicus (Norway rat)
 C;Date: 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change 16-Jul-1999
 C;Accession: A27115
 R;Fung, W.P.; Schreiber, G.
 J. Biol. Chem. 262, 9298-9308, 1987
 A;Title: Structure and expression of the genes for major acute phase alpha-1-protein (th
 A;Reference number: A92653; MUID:87250580; PMID:2439509
 A;Accession: A27115
 A;Status: not compared with conceptual translation
 A;Molecule type: DNA
 A;Residues: 1-315 <FUN>
 C;Genetics:
 A;Gene: MAP1

C;Superfamily: kininogen, cystatin homology
 F;19-65/Domain: cystatin homology (fragment) <CVS>
 Query Match 45.7%; Score 42.5; DB 2; Length 315;
 Best Local Similarity 40.0%; Pred. No. 27;
 Matches 10; Conservative 1; Mismatches 5; Indels 9; Gaps 1;

QY 1 GHXXKLDD-----DLHQGGH 16
 |||||

Db 159 GHKKLDDLKQORDGYNRHPMGH 183

RESULT 10

T27668
 hypothetical protein ZK1053.2 - Caenorhabditis elegans
 C;Species: Caenorhabditis elegans
 C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 04-Mar-2000
 C;Accession: T27668
 R;Kershaw, J.
 submitted to the EMBL Data Library, November 1996
 A;Reference number: Z20402
 A;Accession: T27668
 A;Status: preliminary; translated from GB/EMBL/DDBJ
 A;Molecule type: DNA
 A;Residues: 1-536 <WIL>
 A;Cross-references: EMBL:Z20404; PIDN:CAB04976.1; GSPDB:GN00019; CESP:ZK1053.2
 A;Experimental source: clone ZK1053
 C;Genetics:
 A;Gene: CESP:ZK1053.2
 A;Map position: 1
 A;Introns: 56/3; 88/3; 203/1; 228/2; 383/1; 467/3
 C;Superfamily: Caenorhabditis elegans hypothetical protein ZK1053.2

Query Match 45.7%; Score 42.5; DB 2; Length 536;
 Best Local Similarity 50.0%; Pred. No. 48;
 Matches 8; Conservative 5; Mismatches 0; Indels 3; Gaps 1;

QY 5 KLDDDLHQ---GGHV 17
 :|||::|
 Db 510 RLDDVDVHKPRGGYV 525

RESULT 11

G98017
 hypothetical protein spr1168 [imported] - Streptococcus pneumoniae (strain R6)
 C;Species: Streptococcus pneumoniae
 C;Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 22-Oct-2001
 C;Accession: G98017
 R;Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; B
 y, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M
 y, P.; Sun, P.M.; Winkler, M.E.
 J. Bacteriol. 183, 5709-5717, 2001
 A;Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;
 A;Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
 A;Reference number: A97872; MUID:21429245; PMID:11544234
 A;Accession: G98017
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-125 <KUR>
 A;Cross-references: GB:AE007317; PIDN:AAK99971.1; PID:g15458799; GSPDB:GN00174
 C;Genetics:
 A;Gene: spr1168

Query Match 45.2%; Score 42; DB 2; Length 125;
 Best Local Similarity 50.0%; Pred. No. 12;
 Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 GHXXKLDDDLHQGGH 16
 |||||
 Db 98 GQKELDYQLQNGGH 113

RESULT 12

G97280
 uncharacterized conserved protein, associated with phosphate permease [imported] - Clostri
 C;Species: Clostridium acetobutylicum
 C;Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 14-Sep-2001
 C;Accession: G97280
 R;Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,
 ; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
 J. Bacteriol. 183, 4823-4836, 2001
 A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clo

A:Reference number: A96900; MUID:21359325; PMID:21359325
 A:Accession: G97280
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-210 <KUR>
 A:Cross-references: GB:AB001437; PIDN:AAK81034.1; PID:gl5026159; GSPDB:GN00168
 A:Experimental source: Clostridium acetobutylicum ATCC824
 C:Genetics:
 A:Gene: CAC3094

Query Match 45.2%; Score 42; DB 2; Length 210;
 Best Local Similarity 46.7%; Pred. No. 21;
 Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 2 HKXKLDLLEHGGGH 16
 Db 41 NKLKVKDMEHKGDDH 55

RESULT 13
 T27198
 hypothetical protein Y57A10B.7 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 04-Mar-2000
 C:Accession: T27198
 R:Smyle, R.
 submitted to the EMBL Data Library, October 1998
 A:Reference number: Z20326
 A:Accession: T27198
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-237 <WIL>
 A:Cross-references: EMBL:AL032647; PIDN:CAA21694.1; GSPDB:GN00019; CESP:Y57A10B.7
 A:Experimental source: clone Y57A10B
 C:Genetics:
 A:Gene: CESP:Y57A10B.7
 A:Map position: 1
 A:Introns: 37/3; 109/3; 166/3
 C:Superfamily: Caenorhabditis elegans hypothetical protein Y57A10B.7

Query Match 45.2%; Score 42; DB 2; Length 237;
 Best Local Similarity 50.0%; Pred. No. 24;
 Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 HKXKLDLLEHGGGH 17
 Db 182 HAEKVDNNQGLV 197

RESULT 14
 T23844
 hypothetical protein R01E6.5 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
 C:Accession: T23844
 R:Coles, J.
 submitted to the EMBL Data Library, November 1995
 A:Reference number: Z19807
 A:Accession: T23844
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-245 <WIL>
 A:Cross-references: EMBL:Z68118; PIDN:CAA92185.1; GSPDB:GN00028; CESP:R01E6.5
 A:Experimental source: clone R01E6
 C:Genetics:
 A:Gene: CESP:R01E6.5
 A:Map position: X
 A:Introns: 9/2; 148/1

Query Match 45.2%; Score 42; DB 2; Length 245;
 Best Local Similarity 50.0%; Pred. No. 25;
 Matches 8; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 GHXKLDLLEHGGGH 16
 Db 167 GHKSYYDNEGKYGDH 182

RESULT 15
 K3BOH1
 kininogen, HMW I precursor - bovine
 N:Alternate names: alpha-2-thiol proteinase inhibitor; preprokininogen
 N:Contains: bradykinin (kallidin); kininogen I; kininogen II; prokininogen
 C:Species: Bos primigenius taurus (cattle)
 C>Date: 14-Nov-1983 #sequence_revision 14-Nov-1983 #text_change 22-Jun-1999
 C:Accession: A01281; A91938; A91938; A29559
 R:Kitamura, N.; Takagaki, Y.; Furuto, S.; Tanaka, T.; Nawa, H.; Nakanishi, S.
 Nature 305, 545-549, 1983
 A:Title: A single gene for bovine high molecular weight and low molecular weight kininogen
 A:Reference number: A93317; MUID:84014106; PMID:6571699
 A:Accession: A01281
 A:Molecule type: mRNA
 A:Residues: 1-621 <KIT>
 A:Cross-references: GB:V01491; GB:K01757; NID:g491; PIDN:CAA24735.1; PID:g492
 R:Kato, H.; Nagasawa, S.; Suzuki, T.
 J. Biochem. 67, 313-323, 1970
 A:Title: Studies on the structure of bovine kininogen: cleavages of disulfide bonds and
 A:Reference number: A91923; MUID:70180420; PMID:4986212
 A:Accession: A91923
 A:Molecule type: protein
 A:Residues: 378-393 <KAT>
 R:Han, Y.N.; Komiya, M.; Iwanaga, S.; Suzuki, T.
 J. Biochem. 77, 55-68, 1975
 A:Title: Studies on the primary structure of bovine high-molecular-weight kininogen. Amir
 A:Reference number: A91938; MUID:75170265; PMID:1169237
 A:Accession: A91938
 A:Molecule type: protein
 A:Residues: 458-498 <HAN>
 R:Sueyoshi, T.; Miyata, T.; Hashimoto, N.; Kato, H.; Hayashida, H.; Miyata, T.; Iwanaga, S.
 J. Biol. Chem. 262, 2768-2779, 1987
 A:Title: Bovine high molecular weight kininogen. The amino acid sequence, positions of
 A:Reference number: A92627; MUID:87137530; PMID:3546295
 A:Accession: A29559
 A:Molecule type: protein
 A:Residues: 'Z', 20-123, 'I', 125-127, 'I', 129-378 <SUE>
 R:Lottspeich, F.; Kellermann, J.; Henschel, A.; Foertsch, B.; Muller-Esterl, W.
 Eur. J. Biochem. 152, 307-314, 1985
 A:Title: The amino acid sequence of the light chain of human high-molecular-mass kininogen
 A:Reference number: A91153; MUID:86030270; PMID:4054110
 A:Contents: annotation; bovine cleavage sites; bovine carbohydrate binding sites
 R:Sueyoshi, T.; Miyata, T.; Kato, H.; Iwanaga, S.
 Seikagaku 56, 808, 1984
 A:Title: Disulfide bonds in bovine HMW kininogens.
 A:Reference number: A94300
 A:Contents: annotation; disulfide bonds
 A:Note: article in Japanese
 C:Comment: The HMW kininogen precursor is produced from the same gene as the LMW form as
 C:Comment: kininogen is a cysteine proteinase inhibitor, takes part in initiation of the
 C:Comment: the glycine/histidine/lysine-rich region of HMW kininogen light chain is impor
 C:Comment: bradykinin, released from kininogen by kallikrein, is a potent vasodilator, it
 C:Superfamily: kininogen; cystatin homology
 C:Keywords: alternative splicing; blood coagulation; cysteine proteinase inhibitor; dupli
 F:1-18/Domain: signal sequence #status predicted <SIG>
 F:19-621/Product: HMW kininogen I #status predicted <MAT>
 F:19-379/Product: HMW kininogen I heavy chain #status experimental <HCH>
 F:19-130/Domain: cystatin homology <CY1>
 F:141-252/Domain: cystatin homology <CY2>
 F:263-374/Domain: cystatin homology <CY3>
 F:379-388/Product: lysyl-bradykinin (kallidin II) #status experimental <KBDY>
 F:380-388/Product: bradykinin (kallidin I) #status experimental <BDY>
 F:389-621/Product: HMW kininogen I light chain #status experimental <LCH>
 F:417-488/Region: glycine/histidine/lysine-rich
 F:19/Motif site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experiment
 F:27-591, 82-93, 106-125, 141-144, 205-217, 228-247, 263-266, 327-339, 350-369/Disulfide bonds: #
 F:187, 168, 169, 204/Binding site: carbohydrate (Asn) (covalent) #status experimental

F:136/Binding site: carbohydrate (Thr) (covalent) (partial) #status experimental
F:197/Binding site: carbohydrate (Asn) (covalent) (partial) #status experimental
F:378-379/Cleavage site: Met-Lys (kallikrein) #status experimental
F:382/Modified site: 4-hydroxyproline (Pro) #status predicted
F:388-389/Cleavage site: Arg-Ser (kallikrein) #status experimental
F:398-406,512/Binding site: carbohydrate (Ser) (covalent) #status experimental
F:399,400,520,524,536,548,553,570/Binding site: carbohydrate (Thr) (covalent) #status experimental
F:498-499/Cleavage site: Arg-Thr (kallikrein) #status experimental

Query Match 45.2%; Score 42; DB 1; Length 621;
Best Local Similarity 50.0%; Pred. NO. 68;
Matches 8; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 GHXXKLDLLEHQGH 16
||| |||
Db 444 GHKKHDDQGHGHGH 459

Search completed: March 4, 2004, 13:11:23
Job time : 11.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2004, 12:53:12 ; Search time 6.75 Seconds
(without alignments)
131.140 Million cell updates/sec

Title: SEQ-C

Perfect score: 93

Sequence: 1 ghkxkLDDLEHQGHV 17

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 segs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	91	97.8	644	1	KNG_HUMAN	P01042 homo sapien
2	46	49.5	472	1	NADB_SULSO	Q97ZCS sulfolobus
3	44	47.3	134	1	RETS_HUMAN	P82980 homo sapien
4	43	46.2	683	1	RPC_BPHFC	P08979 bacterioph
5	42	45.2	621	1	KNH1_BOVIN	P01044 bos taurus
6	42	45.2	624	1	STS_MOUSE	P50427 mus musculu
7	42	45.2	1061	1	TRC3_ECOLI	P27189 escherichia
8	42	45.2	1448	1	TRC5_ECOLI	P27190 escherichia
9	42	45.2	1785	1	GLS3_YEAST	Q04952 saccharomyc
10	41.5	44.6	639	1	KNG_RAT	P08934 rattus norv
11	41	44.1	252	1	LIFB_REIME	Q92qds rhizobium m
12	41	44.1	538	1	CP5D_CANVA	P16141 candida mal
13	41	44.1	2332	1	POLG_FMDVA	P03308 f genome po
14	40.5	43.5	661	1	KNG_MOUSE	Q08677 mus musculu
15	40.5	43.5	524	1	YB53_YEAST	P38308 saccharomyc
16	40	43.0	239	1	Y288_MYCGE	P52744 homo sapien
17	40	43.0	414	1	Y288_MYCGE	Q9V577 mycoplasma
18	40	43.0	477	1	TM17_HUMAN	Q9V577 mycoplasma
19	40	43.0	878	1	SVV_METUA	Q58413 methanococc
20	40	43.0	1744	1	TENS_CHICK	Q04205 gallus gall
21	39.5	42.5	121	1	REFA_CLOTE	Q895j7 ciosridium
22	39.5	42.5	1023	1	PIB4_BOVIN	Q07722 bos taurus
23	39	41.9	177	1	RT23_MOUSE	Q8ve22 mus musculu
24	39	41.9	335	1	HYPB_RHOCA	P26410 rhodobacter
25	39	41.9	339	1	OTCC_STEMU	Q8dw19 streptococc
26	39	41.9	355	1	CAPH_STRAU	P39857 scapyllococ
27	39	41.9	436	1	CUS1_YEAST	Q02554 saccharomyc
28	39	41.9	551	1	CEX4_MOUSE	O55187 mus musculu
29	39	41.9	558	1	CEX4_HUMAN	O00257 homo sapien
30	39	41.9	570	1	DLA_ECOLI	P06149 escherichia
31	39	41.9	551	1	BGLR_HUMAN	P08236 homo sapien
32	39	41.9	729	1	KEX1_YEAST	P09620 saccharomyc
33	39	41.9	744	1	REL4_ECOLI	P11585 escherichia

34	39	41.9	1531	1	PMED_CHLTR	O84818 chlamydia t
35	38.5	41.4	544	1	TCPG_HUMAN	P49368 homo sapien
36	38	40.9	85	1	IMW8_ECOLI	P09881 escherichia
37	38	40.9	85	1	IMW8_ECOLI	P13478 escherichia
38	38	40.9	123	1	LCA_CAMDR	P00710 camelus dro
39	38	40.9	573	1	GUBI_MAIZE	P15590 zea mays (m
40	38	40.9	766	1	TKLI_MOUSE	Q8C0V0 mus musculu
41	38	40.9	766	1	TKLI_MOUSE	Q9UKI8 homo sapien
42	38	40.9	818	1	NAH1_PIG	P48762 sus scrofa
43	38	40.9	1270	1	DXH9_HUMAN	Q08211 homo sapien
44	38	40.9	1287	1	DXH9_BOVIN	Q28141 bos taurus
45	37.5	40.3	387	1	VINT_BPP22	P04890 bacterioph

ALIGNMENTS

RESULT 1
KNG_HUMAN
ID KNG_HUMAN STANDARD; PRT; 644 AA.
AC P01042; P01043;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Kininogen precursor (Alpha-2-thiol proteinase inhibitor) [Contains:
DE Bradykinin].
DE KNG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS HMW AND LMW).
RC TISSUE=Liver;
RX MEDLINE=85234582; PubMed=2989293;
RA Takagaki Y., Kitamura N., Nakanishi S.;
RT "Cloning and sequence analysis of cDNAs for human high molecular
RT weight and low molecular weight prekininogens. Primary structures of
RT two human prekininogens.";
RL J. Biol. Chem. 260:8601-8609(1985).
RN [2]
RP GENE STRUCTURE.
RX MEDLINE=85234583; PubMed=2989294;
RA Kitamura N., Kitagawa H., Fukushima D., Takagaki Y., Miyata T.,
RA Nakanishi S.;
RT "Structural organization of the human kininogen gene and a model for
RT its evolution.";
RL J. Biol. Chem. 260:8610-8617(1985).
RN [3]
RP SEQUENCE OF 1-401 FROM N.A.
RX MEDLINE=8512621; PubMed=6441591;
RA Ohkubo I., Kurachi K., Takasawa T., Shiokawa H., Sasaki M.;
RT "Isolation of a human cDNA for alpha 2-thiol proteinase inhibitor and
RT its identity with low molecular weight kininogen.";
RL Biochemistry 23:5691-5697(1984).
RN [4]
RP SEQUENCE OF 379-644.
RX MEDLINE=86030270; PubMed=4054110;
RA Lottspeich F., Kellermann J., Henschen A., Foersts B.,
RA Mueller-Esterl W.;
RT "The amino acid sequence of the light chain of human high-molecular-
RT mass kininogen.";
RL Eur. J. Biochem. 152:307-314(1985).
RN [5]
RP SEQUENCE OF 381-389.
RX MEDLINE=90255622; PubMed=4952632;
RA Pierce J.V.;
RT "Structural features of plasma kinins and kininogens.";
RL Fed. Proc. 27:52-57(1968).
RN [6]
RP DISULFIDE BONDS.
RA Sueyoshi T., Miyata T., Kato H., Iwanaga S.;
RT "Disulfide bonds in bovine HMW kininogens.";

RL Seikagaku 56:808-808(1984).
 RN [7]
 RP CARBOHYDRATE-LINKAGE SITE ASN-294.
 RX MEDLINE=22660472; PubMed=12754519;
 RT Zhang H., Li X.-J., Martin D.B., Aebbersold R.;
 RA "Identification and quantification of N-linked glycoproteins using
 RT hydrazide chemistry, stable isotope labeling and mass spectrometry.";
 RL Nat. Biotechnol. 21:660-666(2003)
 CC -!- FUNCTION: (1) Kininogens are inhibitors of thiol proteases; (2)
 CC HMW-kininogen plays an important role in blood coagulation by
 CC helping to position optimally prekallikrein and factor XI next to
 CC factor XII; (3) HMW-kininogen inhibits the thrombin- and plasmin-
 CC induced aggregation of thrombocytes; (4) the active peptide
 CC bradykinin that is released from HMW-kininogen shows a variety of
 CC physiological effects: (4A) influence in smooth muscle
 CC contraction, (4B) induction of hypotension, (4C) natriuresis and
 CC diuresis, (4D) decrease in blood glucose level, (4E) it is a
 CC mediator of inflammation and causes (4E1) increase in vascular
 CC permeability, (4E2) stimulation of nociceptors (4E3) release of
 CC other mediators of inflammation (e.g. prostaglandins), (4F) it has
 CC a cardioprotective effect (directly via bradykinin action,
 CC indirectly via endothelium-derived relaxing factor action); (5)
 CC LMW-kininogen inhibits the aggregation of thrombocytes; (6) LMW-
 CC kininogen is in contrast to HMW-kininogen not involved in blood
 CC clotting.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=HMW;
 CC IsoId=P01042-1; Sequence=displayed;
 CC Name=LMW;
 CC IsoId=P01042-2; Sequence=VSP_001261, VSP_001262;
 CC -!- TISSUE SPECIFICITY: Plasma.
 CC -!- PTM: Bradykinin is released from kininogen by plasma kallikrein.
 CC -!- SIMILARITY: Contains 3 cystatin-like domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X02566; AAB35497.1; -;
 DR EMBL; M11437; AAB59550.1; JOINED.
 DR EMBL; M11438; AAB59550.1; JOINED.
 DR EMBL; M11521; AAB59550.1; JOINED.
 DR EMBL; M11522; AAB59550.1; JOINED.
 DR EMBL; M11523; AAB59550.1; JOINED.
 DR EMBL; M11524; AAB59550.1; JOINED.
 DR EMBL; M11525; AAB59550.1; JOINED.
 DR EMBL; M11526; AAB59550.1; JOINED.
 DR EMBL; M11527; AAB59550.1; JOINED.
 DR EMBL; M11528; AAB59550.1; JOINED.
 DR EMBL; M11529; AAB59550.1; JOINED.
 DR EMBL; M11437; AAB59551.1; -;
 DR EMBL; M11438; AAB59551.1; JOINED.
 DR EMBL; M11521; AAB59551.1; JOINED.
 DR EMBL; M11522; AAB59551.1; JOINED.
 DR EMBL; M11523; AAB59551.1; JOINED.
 DR EMBL; M11524; AAB59551.1; JOINED.
 DR EMBL; M11525; AAB59551.1; JOINED.
 DR EMBL; M11526; AAB59551.1; JOINED.
 DR EMBL; M11527; AAB59551.1; JOINED.
 DR EMBL; M11528; AAB59551.1; JOINED.
 DR PIR; A01279; KGHUHL.
 DR PIR; A01280; KGHULL.
 DR SWISS-2DPAGE; P01042; HUMAN.
 DR Genew; HGNC:6383; KNG.
 DR MIM; 228960;
 DR GO; GO:0007596; P:blood coagulation; NAS.
 DR GO; GO:0030146; P:diuresis; NAS.
 DR GO; GO:0006954; P:inflammatory response; NAS.

DR GO; GO:0030147; P:natriuresis; NAS.
 DR GO; GO:0006939; P:smooth muscle contraction; NAS.
 DR InterPro; IPR00010; Cystatin.
 DR InterPro; IPR002395; Kininogen.
 DR Pfam; PF00031; cystatin; 3.
 DR PRINTS; PR00334; KININOGEN.
 DR SMART; SM00043; CY; 3.
 DR PROSITE; PS00287; CYSTATIN; 2.
 KW Glycoprotein; Plasma; Repeat; Thiol protease inhibitor; Vasodilator;
 KW Bradykinin; Blood coagulation; Inflammatory response; Signal;
 KW Alternative splicing; Pyrrolidone carboxylic acid.
 FT SIGNAL 1 18
 FT CHAIN 19 644
 FT CHAIN 19 380
 FT PEPTIDE 381 389
 FT CHAIN 390 644
 FT CHAIN 19 136
 FT DOMAIN 137 258
 FT DOMAIN 259 380
 FT DOMAIN 420 510
 FT REPEAT 420 449
 FT REPEAT 450 479
 FT REPEAT 480 510
 FT MOD RES 19 19
 FT DISULFID 28 614
 FT DISULFID 83 94
 FT DISULFID 107 126
 FT DISULFID 142 145
 FT DISULFID 206 218
 FT DISULFID 229 248
 FT DISULFID 264 267
 FT DISULFID 328 340
 FT DISULFID 351 370
 FT CARBOHYD 48 48
 FT CARBOHYD 169 169
 FT CARBOHYD 205 205
 FT CARBOHYD 294 294
 FT CARBOHYD 401 401
 FT CARBOHYD 533 533
 FT CARBOHYD 542 542
 FT CARBOHYD 546 546
 FT CARBOHYD 557 557
 FT CARBOHYD 571 571
 FT CARBOHYD 577 577
 FT CARBOHYD 593 593
 FT CARBOHYD 628 628
 FT VARSPLIC 402 427
 FT VARSPLIC 428 644
 FT CONFLICT 593 593
 FT SEQUENCE 644 AA; 71945 NW; 3132B4CBARFBB7E CRC64;
 Query Match 97.8%; Score 91; DB 1; Length 644;
 Best Local Similarity 94.1%; Pred. No. 4.2e-07;
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 GHKXKLDLLEHOGGHV 17
 DB 474 GHKXKLDLLEHOGGHV 490
 RESULT 2
 NAME SULSO
 ID NAME SULSO STANDARD; PRT; 472 AA.
 AC Q97ZCS;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE L-aspartate oxidase (EC 1.4.3.16) (LASPO) (Quinolinic synthetase B).
 GN NADB OR SSO0997.

```

OS Sulfolobus solfataricus.
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC Sulfolobus.
OX NCBI_TaxID=2287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35092 / DSM 1617 / P2;
RX MEDLINE=21332296; PubMed=11427726;
RA She Q., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,
RA Aways M.J., Chan-Weller C.C.-Y., Clausen I.G., Curtis B.A.,
RA De Moors G., Erasus G., Fletcher C., Gordon P.M.K.,
RA Heikamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,
RA Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,
RA Charlebois R.L., Doolittle W.F., Duquet M., Gaasterland T.,
RA Garrett R.A., Regan M.A., Sensen C.W., Van der Oost J.;
RT "The complete genome of the crenarchaeon Sulfolobus solfataricus P2.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).
CC -!- FUNCTION: Catalyzes the oxidation of L-aspartate to
CC L-lysine.
CC -!- CATALYTIC ACTIVITY: L-aspartate + H(2)O + O(2) = oxaloacetate +
CC NH(3) + H(2)O(2).
CC -!- COFACTOR: FAD.
CC -!- PATHWAY: NAD biosynthesis; aspartate to NAMN; first step.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: BELONGS TO THE FAD-DEPENDENT OXIDOREDUCTASE FAMILY 2.
CC NADB SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; A3006719; AAK41268.1; AUT INIT.
DR InterPro; IPR003953; FAD bind2.
DR InterPro; IPR0041327; FAD_PVT redox.
DR InterPro; IPR004112; Succ_DH flav_C.
DR Pfam; PF00890; FAD binding_2; 2.
DR Pfam; PF02910; succ_DH flav_C; 1.
DR PRINTS; PR00368; FADPNR.
KW Pyridine nucleotide biosynthesis; Oxidoreductase; Flavoprotein; FAD;
KW Complete proteome.
FT NP BIND 3 17 FAD (AMP PART) (POTENTIAL).
FT ACT SITE 211 211 BY SIMILARITY.
FT ACT SITE 229 229 BY SIMILARITY.
SQ SEQUENCE 472 AA; 52644 MW; DABAB606463C353D CRC64;
Query Match 49.5%; Score 46; DB 1; Length 472;
Best Local Similarity 66.7%; Pred. No. 5.8;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 5 KLDDLEHQGH 16
Db | ||| : |||
94 KFTDDLRLGGH 105
RESULT 3
IDS HUMAN
RETS HUMAN STANDARD; PRT; 134 AA.
AC P82980;
DT 16-OCT-2001 (Rel. 40, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Retinol-binding protein III, cellular (CRBP-III) (HRBPiso).
GN RBP5.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS), FUNCTION,

```

```

RP AND TISSUE SPECIFICITY.
RX MEDLINE=21173623; PubMed=11274389;
RA Polli C., Calderone V., Ottonello S., Bolchi A., Zanotti G.,
RA Stoppini M., Berni R.;
RT "Identification, retinoid binding and X-ray analysis of a human
RT retinol-binding protein.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:3710-3715(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Li Y., Wu T., Xu S., Ren S., Chen Z., Han Z.;
RT "A novel gene expressed in human liver non-tumor tissue.";
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Whiting J., Helton E., Kettman M., Maman A., Rodrigues S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grinstead J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywicki M.I., Skalska U., Smalish D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: Intracellular transport of retinol.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- TISSUE SPECIFICITY: HIGHER EXPRESSION IN ADULT KIDNEY AND LIVER
CC AND TO A LESSER EXTENT IN ADULT AND FETAL SPLEEN, ADULT LYMPH
CC NODES AND APPENDIX, AND FETAL LIVER AND KIDNEY.
CC -!- SIMILARITY: Belongs to the fatty-acid binding protein (FABP)
CC family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AY007436; AAG09617.1; -.
DR EMBL; AF212239; AAK14925.1; -.
DR EMBL; BC029355; AAK29355.1; -.
DR PDB; 1GGL; 04-APR-01.
DR Genew; HGNC:15847; RBP5.
DR GO; GO:0005737; C:cytoplasm; NAS.
DR GO; GO:0005478; F:intracellular transporter activity; TAS.
DR GO; GO:0005501; F:retinoid binding; TAS.
DR GO; GO:0006810; P:transport; NAS.
DR InterPro; IPR000463; Fatty acid BP.
DR InterPro; IPR000566; Lipocalin_cytfabp.
DR Pfam; PF00061; lipocalin; 1.
DR PRINTS; PR00178; FATTYACIDBP.
DR PROSITE; PS00214; FABP; FALSE NEG.
KW Vitamin A; Retinol-binding; Transport; 3D-structure.
FT INIT MET 0
SQ SEQUENCE 134 AA; 15800 MW; E6E959DC9333B124 CRC64;
Query Match 47.3%; Score 44; DB 1; Length 134;
Best Local Similarity 53.8%; Pred. NO. 3.2;

```

Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 5 KLDDLEHQGHV 17
| : ||||| :
Db 37 KPDKEIEHQNHM 49

RESULT 4

RPC_BPPHC STANDARD; PRT; 683 AA.
ID -RPC_BPPHC
AC P08979;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Repressor protein C.
GN C.
OS Bacteriophage phi-C31.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae;
OC Lambda-like viruses.
OX NCBI_TaxID=10719;
RN [1]
RP SEQUENCE FROM N.A.
RC SPRAIN=Norwich;
RX MEDLINE=89039715; PubMed=3185504;
RA Sinclair R.B., Bibb M.J.;
RT "The repressor gene (c) of the Streptomyces temperate phage phi c31:
RT nucleotide sequence, analysis and functional cloning.";
RL Mol. Gen. Genet. 213:269-277(1988).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=94374705; PubMed=8088546;
RA Hartley N.M., Murphy G.O., Bruton C.J., Chater K.F.;
RT "Sequence of the essential early region of phi C31, a temperate phage
RT of Streptomyces spp. with unusual features in its lytic
RT development.";
RL Gene 147:29-40(1994).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X12865; CAA31345.1; -;
DR EMBL; X76288; CAA53911.1; -;
DR PIR; S01433; S01433.
DR KW Transcription regulation; Repressor; DNA-binding.
SQ SEQUENCE 683 AA; 74077 MW; B02379D204F37D1B CRC64;

Query Match 46.2%; Score 43; DB 1; Length 683;
Best Local Similarity 70.0%; Pred. No. 26;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 7 DDLEHQGH 16
| : ||||| :
Db 474 DDVERQGAH 483

RESULT 5

KNH1_BOVIN STANDARD; PRT; 621 AA.
ID -KNH1_BOVIN
AC P01044;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Kininogen, HMW I precursor (Thiol proteinase inhibitor) [Contains:
DE Bradykinin].
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euteleostomi; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.

OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=84014106; PubMed=6571699;
RA Kitamura N., Takagaki Y., Furuto S., Tanaka T., Nawa H., Nakanishi S.;
RT "A single gene for bovine high molecular weight and low molecular
RT weight kininogens.";
RL Nature 305:545-549(1983).
RN [2]
RP SEQUENCE OF 19-378.
RX MEDLINE=87137530; PubMed=3546295;
RA Sueyoshi T., Miyata T., Hashimoto N., Kato H., Hayashida H.,
RA Miyata T., Iwanaga S.;
RT "Bovine high molecular weight kininogen. The amino acid sequence,
RT positions of carbohydrate chains and disulfide bridges in the heavy
RT chain portion.";
RL J. Biol. Chem. 262:2768-2779(1987).
RN [3]
RP SEQUENCE OF 378-393.
RX MEDLINE=70180420; PubMed=4986212;
RA Kato H., Nagasawa S., Suzuki T.;
RT "Studies on the structure of bovine kininogen: cleavages of disulfide
RT bonds and of methionyl bonds in kininogen-II.";
RL J. Biochem. 67:313-323(1970).
RN [4]
RP SEQUENCE OF 458-498.
RX MEDLINE=75170265; PubMed=1169237;
RA Han Y.N., Komiya M., Iwanaga S., Suzuki T.;
RT "Studies on the primary structure of bovine high-molecular-weight
RT kininogen. Amino acid sequence of a fragment ('histidine-rich
RT peptide') released by plasma kallikrein.";
RL J. Biochem. 77:55-68(1975).
CC
CC -!- FUNCTION: (1) Kininogens are inhibitors of thiol proteases; (2)
CC HMW-kininogen plays an important role in blood coagulation by
CC helping to position optimally prekallikrein and factor XI next to
CC factor XII; (3) HMW-kininogen inhibits the thrombin- and plasmin-
CC induced aggregation of thrombocytes; (4) the active peptide
CC bradykinin that is released from HMW-kininogen shows a variety of
CC physiological effects: (4A) influence in smooth muscle
CC contraction, (4B) induction of hypotension, (4C) natriuresis and
CC diuresis, (4D) decrease in blood glucose level, (4E) it is a
CC mediator of inflammation and causes (4E1) increase in vascular
CC permeability, (4E2) stimulation of nociceptors (4E3) release of
CC other mediators of inflammation (e.g. prostaglandins), (4F) it has
CC a cardioprotective effect (directly via bradykinin action,
CC indirectly via endothelium-derived relaxing factor action).
CC
CC -!- SUBCELLULAR LOCATION: Extracellular.
CC
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=HMW I;
CC IsoID=P01044-1; Sequence=Displayed;
CC Name=LMW I;
CC IsoID=P01046-1; Sequence=External;
CC
CC -!- TISSUE SPECIFICITY: Plasma.
CC
CC -!- PM: Bradykinin is released from kininogen by plasma kallikrein.
CC
CC -!- SIMILARITY: Contains 3 cystatin-like domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; V01491; CAA24735.1; -;
DR PIR; A01281; KBOH1.
DR InterPro; IPR000010; Cystatin.
DR InterPro; IPR002395; Kininogen.
DR Pfam; PF00031; cystatin; 3.
DR PRINTS; PR00334; KININOGEN.
DR SMART; SM00043; CY; 3.
DR PROSITE; PS00287; CYSTATIN; 2.

KW Glycoprotein; Plasma; Repeat; Vasodilator; Alternative splicing;
 KW Thiol protease inhibitor; Bradykinin; Blood coagulation;
 KW Inflammatory response; Signal; Pyrrolidone carboxylic acid.
 FT SIGNAL 1 18
 FT CHAIN 19 621 KININOGEN, HMW I.
 FT CHAIN 19 378 HEAVY CHAIN.
 FT PEPTIDE 380 388 BRADYKININ.
 FT CHAIN 389 621 LIGHT CHAIN.
 FT DOMAIN 19 135 CYSTATIN-LIKE 1.
 FT DOMAIN 136 257 CYSTATIN-LIKE 2.
 FT DOMAIN 258 378 CYSTATIN-LIKE 3.
 FT MOD_RES 19 19 PYRROLIDONE CARBOXYLIC ACID.
 FT CARBOHYD 87 87 N-LINKED (GLCNAC. . .).
 FT CARBOHYD 136 136 O-LINKED (PARTIAL).
 FT CARBOHYD 168 168 N-LINKED (GLCNAC. . .) (OR 169).
 FT CARBOHYD 197 197 N-LINKED (GLCNAC. . .) (PARTIAL).
 FT CARBOHYD 204 204 N-LINKED (GLCNAC. . .).
 FT DISULFID 27 591 INTERCHAIN.
 FT DISULFID 82 93
 FT DISULFID 106 125
 FT DISULFID 141 144
 FT DISULFID 205 217
 FT DISULFID 228 247
 FT DISULFID 263 286
 FT DISULFID 327 339
 FT DISULFID 350 369
 SQ SEQUENCE 621 AA; 68890 MW; D16850BEFE3C55CD CRC64;
 Query Match 45.2%; Score 42; DB 1; Length 621;
 Best Local Similarity 50.0%; Pred. No. 34;
 Matches 8; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
 QY 1 GHKXKLDLLEHGQH 16
 Db 444 GHKXKLDLLEHGQH 459
 RESULT 6
 STS_MOUSE
 ID STS_MOUSE STANDARD; PRT; 624 AA.
 AC P50427;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Steryl-sulfatase precursor (EC 3.1.6.2) (steroid sulfatase) (Steryl-sulfate sulfohydrolase) (Arylsulfatase C) (ASC).
 GN SNS.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BAIE/C;
 RX MEDLINE=96241577; PubMed=8673109;
 RA Salido E.C., Li X.M., Yen P.H., Martin N., Mohandas T.K.,
 RA Shapiro L.J.;
 RT "Cloning and expression of the mouse pseudoautosomal steroid sulphatase gene (Sts).";
 RL Nat. Genet. 13:83-86(1996).
 CC -1- FUNCTION: Conversion of sulfated steroid precursors to estrogens during pregnancy.
 CC -1- CATALYTIC ACTIVITY: 3-beta-hydroxyandrost-5-en-17-one 3-sulfate + H(2)O = 3-beta-hydroxyandrost-5-en-17-one + sulfate.
 CC -1- SUBUNIT: Homodimer.
 CC -1- SUBCELLULAR LOCATION: Microsomal membrane. The sequence shows several membrane-spanning domains that could serve to anchor the protein in the microsomal membrane.
 CC -1- SIMILARITY: Belongs to the sulfatase family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/)

CC use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U37545; AAB09308.1; -.
 DR HSP; P15848; 1FSU.
 DR MGD; MGI:98438; Sts.
 DR InterPro; IPR000917; Sulfatase.
 DR Pfam; PF0884; Sulfatase, 1.
 DR PROSITE; PS00523; SULFATASE 1; 1.
 DR PROSITE; PS00149; SULFATASE 2; 1.
 KW Hydrolase; Microsome; Transmembrane; Glycoprotein; Steroid metabolism;
 KW Pregnancy; Signal.
 FT SIGNAL 1 22 POTENTIAL.
 FT CHAIN 23 624 STERYL-SULFATASE.
 FT MOD_RES 83 83 2-AMINO-3-OXOPROPIONIC ACID (BY SIMILARITY).
 FT ACT_SITE 144 144 POTENTIAL.
 FT TRANSMEM 185 205 POTENTIAL.
 FT TRANSMEM 214 234 POTENTIAL.
 FT CARBOHYD 55 55 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 465 465 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 624 AA; 66590 MW; 025C0CF8659D0A9E CRC64;
 Query Match 45.2%; Score 42; DB 1; Length 624;
 Best Local Similarity 50.0%; Pred. No. 35;
 Matches 8; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
 QY 1 GHKXKLDLLEHGQH 16
 Db 599 GHKXKLDLLEHGQH 614
 RESULT 7
 TRC4_ECOLI
 ID TRC4_ECOLI STANDARD; PRT; 1061 AA.
 AC P27189; P27184;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE DNA primase trc (EC 2.7.7.-) (Replication primase).
 GN TRAC.
 OS Escherichia coli.
 OG Plasmid Incp-alpha RP4.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-14 AND 316-333.
 RC STRAIN=HB101;
 RX MEDLINE=92297959; PubMed=1818755;
 RA Mele L., Strack B., Kruff V., Lanka E.;
 RT "Gene organization and nucleotide sequence of the primase region of IncP plasmids RP4 and R751.";
 RL DNA Seq. 2:145-162(1991).
 CC -1- FUNCTION: REQUIRED FOR AUTONOMOUS REPLICATION IN ESCHERICHIA COLI. TRANSFERRED INTO THE RECIPIENT CELL DURING BACTERIAL CONJUGATION. CATALYZES THE SYNTHESIS OF SHORT OLIGORIBONUCLEOTIDE PRIMERS WITH CPA. . . OR PCPA. . . AT THEIR 5'-TERMINI ON A SINGLE STRANDED TEMPLATE DNA.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative initiation;
 CC Comment=2 isoforms, trc-1 (shown here) and trc-2, are produced by alternative initiation;
 CC -1- SIMILARITY: TO PLASMID R751 TRAC.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/)


```

or send an email to license@isb-sib.ch).
-----
CC EMBL; X59793; CAA42455.1; -
CC DR EMBL; X59793; CAA42455.1; -
CC DR PIR; S37667; S37667.
CC DR InterPro; IPR006171; Toprim_dom.
CC DR InterPro; IPR006154; Toprim_sub.
CC DR Pfam; PF01751; Toprim; 1.
CC DE SMART; SM00493; TOPRIM; 1.
CC KW DNA replication; Transferase; DNA-directed RNA polymerase; Plasmid;
CC FT CHAIN 1 1061
CC FT CHAIN 316 1061
CC FT INIT MET 316 316
CC SQ SEQUENCE 1061 AA; 116722 MW; ABC344D2811B931 CRC64;
Query Match 45.2%; Score 42; DB 1; Length 1061;
Best local similarity 53.3%; Pred. No. 61;
Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;
QY 1 GHKXKLDLDLEHGG 15
DB 214 GHPSLRDLRLAHPFG 228
RESULT 8
TRCS_ECOLI
ID TRCS_ECOLI STANDARD; PRT; 1448 AA.
AC P27190; P27195; P27191;
DT 01-NOV-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE DNA primase traC (EC 2.7.7.-) (Replication primase).
GN TRAC
OS Escherichia coli.
OG Plasmid IncP-beta R751.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OC NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 2-11; 219-234 AND 702-714.
RC STRAIN=HB101;
RX MEDLINE=92297959; PubMed=1818755;
RA Miele L., Strack B., Krut V., Lanka E.;
RT "Gene organization and nucleotide sequence of the primase region of
RT IncP plasmids RP4 and R751.";
RL DNA Seq. 2:145-162 (1991).
RL [2]
RN SEQUENCE FROM N.A.
RA Thomas C.M.;
RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: REQUIRED FOR AUTONOMOUS REPLICATION IN ESCHERICHIA COLI.
CC TRANSFERRED INTO THE RECIPIENT CELL DURING BACTERIAL CONJUGATION.
CC CATALYZES THE SYNTHESIS OF SHORT OLIGORIBONUCLEOTIDE PRIMERS WITH
CC CPA OR PCPA AT THEIR 5'-TERMINI ON A SINGLE STRANDED TEMPLATE DNA.
CC -!- ALTERNATIVE PRODUCTS:
CC Event-alternative initiation;
CC Comment-3 isoforms, traC-2 (shown here), traC-3 and traC-4, are
CC produced by alternative initiation;
CC -!- SIMILARITY: TO PLASMID INCP-ALPHA RP4 TRAC.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; X59794; CAA42460.1; -
CC DR EMBL; U67194; AAC64468.1; -
CC DR PIR; S37669; S37669.
CC DR InterPro; IPR006171; Toprim_dom.

```

```

DR InterPro; IPR006154; Toprim_sub.
DR SMART; SM00493; TOPRIM; 1.
DR KW DNA replication; Transferase; DNA-directed RNA polymerase; Plasmid;
CC FT CHAIN 2 1448
CC FT CHAIN 219 1448
CC FT CHAIN 702 1448
CC FT INIT MET 1 1
CC FT INIT MET 219 219
CC FT INIT MET 702 702
CC SQ SEQUENCE 1448 AA; 158950 MW; 1C0EA67CD0D4ED4 CRC64;
Query Match 45.2%; Score 42; DB 1; Length 1448;
Best local similarity 53.3%; Pred. No. 84;
Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;
QY 1 GHKXKLDLDLEHGG 15
DB 431 GHASLRDLRLAHPFG 445
RESULT 9
GLS3_YEAST
ID GLS3_YEAST STANDARD; PRT; 1785 AA.
AC Q04952;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE 1,3-beta-glucan synthase component FKS3 (EC 2.4.1.34) (1,3-beta-D-
DE glucan-UDP glucosyltransferase).
GN FKS3 OR YMR306W OR YMR952.08.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetaceae; Saccharomycetes.
OC NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c / AB972;
RX MEDLINE=9713268; PubMed=9169872;
RA Bowman S., Churcher C.M., Badcock K., Brown D., Chillingworth T.,
RA Connor R., Dedman K., Devlin K., Gentles S., Hamlin N., Hunt S.,
RA Jagels K., Lye G., Moule S., Odell C., Pearson D., Raeburn M.A.,
RA Rice P., Skellon J., Walsh S., Whitehead S., Barrell B.G.;
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome
RT XIII.";
RL Mature 387:90-93 (1997).
CC -!- CATALYTIC ACTIVITY: UDP-glucose + {(1,3)-beta-D-glucosyl}(N) = UDP
CC + {(1,3)-beta-D-glucosyl}(N+1).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -!- SIMILARITY: Belongs to the glycosyltransferase family 48.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; Z49212; CAA89139.1; -
CC DR PIR; S53976; S53976.
CC DR GerMOnline; 142985; -.
CC DR SGD; S0004923; FKS3.
CC DR InterPro; IPR003440; Glyco_trans_48.
CC DR Pfam; PF02364; Glucan synthase; 1.
CC KW Transmembrane; Transferase; Glycosyltransferase.
CC FT TRANSMEM 337 357
CC FT TRANSMEM 375 395
CC FT TRANSMEM 415 435
CC FT TRANSMEM 444 464
CC FT TRANSMEM 508 528
CC FT TRANSMEM 547 567

```

```

FT TRANSMEM 572 592 POTENTIAL.
FT TRANSMEM 712 732 POTENTIAL.
FT TRANSMEM 1215 1235 POTENTIAL.
FT TRANSMEM 1268 1288 POTENTIAL.
FT TRANSMEM 1303 1323 POTENTIAL.
FT TRANSMEM 1370 1390 POTENTIAL.
FT TRANSMEM 1394 1414 POTENTIAL.
FT TRANSMEM 1475 1495 POTENTIAL.
FT TRANSMEM 1514 1534 POTENTIAL.
FT TRANSMEM 1549 1569 POTENTIAL.
FT TRANSMEM 1585 1605 POTENTIAL.
FT TRANSMEM 1655 1675 POTENTIAL.
FT TRANSMEM 1713 1733 POTENTIAL.
SQ SEQUENCE 1785 AA; 207482 MW; 3475446DA46C6120 CRC64;

Query Match 45.2%; Score 42; DB 1; Length 1785;
Best Local Similarity 43.8%; Pred. No. 1.le+02;
Matches 7; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

Qy 1 GHXKKLDDLEHQGGH 16
Db 1443 GYKSKTVDLSEDSGH 1463

RESULT 10
KNG RAT STANDARD; PRT; 639 AA.
ID_KNG RAT
AC P08934; P08933;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Kinnogen precursor [Contains: Bradykinin].
KNG.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RX MEDLINE=87137443; PubMed=3029068;
RA Kitagawa H., Kitamura N., Hayashida H., Miyata T., Nakanishi S.;
RT "Differing expression patterns and evolution of the rat kinnogen
RT gene family."
RL J. Biol. Chem. 262:21190-21198(1987).
RN [2]
RX SEQUENCE FROM N.A. (ISOFORM LMW).
RA Furuto-Kato S., Matsumoto A., Kitamura N., Nakanishi S.;
RT "Primary structures of the mRNAs encoding the rat precursors for
RT bradykinin and T-kinin. Structural relationship of kinnogens with
RT major acute phase protein and alpha 1-cysteine proteinase
RT inhibitor."
RL J. Biol. Chem. 260:12054-12059(1985).
RN [3]
RX SEQUENCE OF 1-65 FROM N.A.
RA STRAIN=Buffalo;
RX MEDLINE=87250580; PubMed=2439509;
RA Fung W.-P., Schreiber G.;
RT "Structure and expression of the genes for major acute phase alpha 1-
RT protein (thiostatin) and kinnogen in the rat."
RL J. Biol. Chem. 262:9298-9308(1987).
RN [4]
RX SEQUENCE OF 1-41 FROM N.A.
RA STRAIN=Wistar; TISSUE=Liver;
RX MEDLINE=87137465; PubMed=3818598;
RA Kageyama R., Kitamura N., Okubo H., Nakanishi S.;
RT "Differing utilization of homologous transcription initiation sites
RT of rat K and T kinnogen genes under inflammation condition."
RL J. Biol. Chem. 262:2345-2351(1987).
RN [5]
RX FUNCTION: (1) Kinnogens are inhibitors of thiol proteases; (2)
RX HMW-kinnogen plays an important role in blood coagulation by
RX helping to position optimally prekallikrein and factor XI next to
RX factor XII; (3) HMW-kinnogen inhibits the thrombin- and plasmin-

```

```

CC induced aggregation of thrombocytes; (4) the active peptide
CC bradykinin that is released from HMW-kinnogen shows a variety of
CC physiological effects: (4A) influence in smooth muscle
CC contraction, (4B) induction of hypotension, (4C) natriuresis and
CC diuresis, (4D) decrease in blood glucose level, (4E) it is a
CC mediator of inflammation and causes (4E1) increase in vascular
CC permeability, (4E2) stimulation of nociceptors (4E3) release of
CC other mediators of inflammation (e.g. prostaglandins), (4F) it has
CC a cardioprotective effect (directly via bradykinin action,
CC indirectly via endothelium-derived relaxing factor action); (5)
CC LMW-kinnogen inhibits the aggregation of thrombocytes; (6) LMW-
CC kinnogen is in contrast to HMW-kinnogen not involved in blood
CC clotting.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=HMW;
CC IsoId=P08934-1; Sequence=Displayed;
CC Name=LMW;
CC IsoId=P08934-2; Sequence=VSP_001265, VSP_001266;
CC -!- TISSUE SPECIFICITY: Plasma.
CC -!- PTM: Bradykinin is released from kinnogen by plasma kallikrein.
CC -!- MISCELLANEOUS: Rats express four types of kinnogens: the
CC classical HMW/LMW kinnogens and two additional LMW-like
CC kinnogens: T-I and T-II.
CC -!- SIMILARITY: Contains 3 cystatin-like domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; L29428; AAA41486.1; -
CC EMBL; M11884; AAA41487.1; -
CC EMBL; M14369; AAA41484.1; -
CC EMBL; M14369; AAA41485.1; ALT_SEQ.
CC EMBL; M16455; AAA41482.1; -
CC PIR; A25486; A25486.
CC PIR; A28055; A28055.
CC InterPro; IPR000010; Cystatin.
CC InterPro; IPR002395; Kinnogen.
CC Pfam; PF00031; cystatin; 3.
CC PRINTS; PR00334; KININOGEN.
CC SMART; SM00043; CY; 3.
CC PROSITE; PS00287; CYSTATIN; 2.
KW Glycoprotein; Plasma; Repeat; Thiol protease inhibitor; Vasodilator;
KW Bradykinin; Blood coagulation; Inflammatory response; Signal;
KW Alternative splicing; Multigene family.
FT SIGNAL 1 18
FT CHAIN 19 639 KININOGEN.
FT CHAIN 19 380 KININOGEN HEAVY CHAIN.
FT PEPTIDE 381 389 BRADYKININ.
FT CHAIN 390 639 KININOGEN LIGHT CHAIN.
FT DOMAIN 19 136 CYSTATIN-LIKE 1.
FT DOMAIN 137 258 CYSTATIN-LIKE 2.
FT DOMAIN 259 380 CYSTATIN-LIKE 3.
FT DOMAIN 439 514 HIS-RICH.
FT DISULFID 28 609 INTERCHAIN (BY SIMILARITY).
FT DISULFID 83 94 BY SIMILARITY.
FT DISULFID 107 126 BY SIMILARITY.
FT DISULFID 142 145 BY SIMILARITY.
FT DISULFID 206 218 BY SIMILARITY.
FT DISULFID 229 248 BY SIMILARITY.
FT DISULFID 264 267 BY SIMILARITY.
FT DISULFID 328 340 BY SIMILARITY.
FT DISULFID 351 370 BY SIMILARITY.
FT CARBOHYD 82 82 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 127 127 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 169 169 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 205 205 N-LINKED (GLCNAC. . .) (POTENTIAL).

```

FT CARBOHYD 294 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 529 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARSPLIC 402 VSPYIARVQEDRDGNEQPIHGWLHAKQ -> RLINS
 FT CEVGRLLKAGAPAPERAQEAETVP (in isoform
 FT LMW).
 FT /FTid=VSP 001265.
 FT Missing (in isoform LMW).
 FT /FTid=VSP 001266.
 FT CONFLICT 61 E -> K (IN REF. 2).
 FT SEQUENCE 639 AA; 70933 MW; D3172DF94FF56AF5 CRC64;
 Query Match 44.6%; Score 41.5; DB 1; Length 639;
 Best Local Similarity 40.0%; Pred. No. 43;
 Matches 10; Conservative 1; Mismatches 5; Indels 9; Gaps 1;
 QY 1 GHKXKLDLDD-----DLEHGGH 16
 DB 461 GHQLKLDLQKQREDGYDRHFGVGH 485
 RESULT 11
 LIPB RHIME
 ID LIPB_RHIME STANDARD; PRT; 252 AA.
 AC Q92QD5;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Lipote-protein ligase B (EC 6.6.1.1) (Lipote biosynthesis protein
 DE B).
 GN LIPB OR R01396 OR SMC01268.
 OS Rhizobium meliloti (Sinorhizobium meliloti).
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.
 OX NCBI_TaxID=382;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=1021;
 RX MEDLINE=21396507; PubMed=11481430;
 RA Capela D., Barloy-Hubler F., Guzy J., Bothe G., Ampe F., Batut J.,
 RA Bolcard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,
 RA Godrie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masuy D.,
 RA Pohl T., Portetelle D., Puehler A., Purnelle B., Ramsperger U.,
 RA Renard C., Thebaud P., Vandenbol M., Weidner S., Galibert F.,
 RA "Analysis of the chromosome sequence of the legume symbiont
 RT Sinorhizobium meliloti strain 1021.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).
 CC -!- FUNCTION: Involved in the attachment of lipoyl groups to proteins,
 CC by creating an amide linkage that joins the free carboxyl group of
 CC lipoyl acid to the epsilon-amino group of a specific lysine
 CC residue in lipoylated proteins (by similarity).
 CC -!- PATHWAY: Lipote biosynthesis.
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -!- SIMILARITY: Belongs to the lipB family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AL591787; CAC45975.1; -.
 CC HAMAP; MF 00013; -; 1.
 DR InterPro; IPR004143; BPL LipA LipB.
 DR InterPro; IPR000544; Lipotease E.
 DR Pfam; PF03039; BPL LipA LipB; 1.
 DR ProDom; PD006086; Lipotease B; 1.
 DR TIGRFAMs; TIGR00214; lipB; 1.
 DR PROSITE; PS01313; LipB; 1.
 KW Ligase; Complete proteome.
 SEQUENCE 252 AA; 28123 MW; 5B2AA9D6012B899D CRC64;

Query Match 44.1%; Score 41; DB 1; Length 252;
 Best Local Similarity 52.9%; Pred. No. 19;
 Matches 9; Conservative 1; Mismatches 7; Indels 0; Gaps 0;
 QY 1 GHKXKLDLDDLEHGGH 17
 DB 192 GLSLNVDPDLHFGGIV 208
 RESULT 12
 CP5D CANNA
 ID CP5D_CANNA STANDARD; PRT; 538 AA.
 AC P16141;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Cytochrome P450 52A4 (EC 1.14.14.-) (CYPLI1A4) (Alkane-inducible P450-
 DE ALK3-A) (P450-CM2).
 GN CYP52A4.
 OS Candida maltosa (Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
 OX NCBI_TaxID=5479;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=EHISD;
 RX MEDLINE=89286595; PubMed=2735924;
 RA Schunck W.-H., Kaergel E., Gross B., Wiedmann B., Mauerberger S.,
 RA Koepke K., Kiessling U., Strauss M., Gaestel M., Mueller H.-G.;
 RT "Molecular cloning and characterization of the primary structure of
 RT the alkane hydroxylating cytochrome P-450 from the yeast Candida
 RT maltosa.";
 RL Biochem. Biophys. Res. Commun. 161:843-850(1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91229697; PubMed=2039569;
 RA Ohkuma M., Tanimoto T., Yano K., Takagi M.;
 RT "CYP52 (cytochrome P450alk) multigene family in Candida maltosa:
 RT molecular cloning and nucleotide sequence of the two tandemly
 RT arranged genes.";
 RL DNA Cell Biol. 10:271-282(1991).
 RN [3]
 RP CHARACTERIZATION.
 RX MEDLINE=96311366; PubMed=8713123;
 RA Zimmer T., Ohkuma M., Ohta A., Takagi M., Schunck W.H.;
 RT "The CYP52 multigene family of Candida maltosa encodes functionally
 RT diverse n-alkane-inducible cytochromes P450.";
 RL Biochem. Biophys. Res. Commun. 224:784-789(1996).
 CC -!- FUNCTION: Together with an NADPH cytochrome P450 the enzyme system
 CC catalyzes the terminal hydroxylation as the first step in the
 CC assimilation of alkanes and fatty acids.
 CC -!- CATALYTIC ACTIVITY: Hydroxylation of N-alkanes at the terminal
 CC position.
 CC -!- INDUCTION: By alkanes.
 CC -!- SIMILARITY: Belongs to the cytochrome P450 family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X51932; CAA36198.1; -.
 DR PIR; S08668; O4CRA4.
 DR HSP; P14779; 1JPZ.
 DR InterPro; IPR001128; Cytochrome_P450.
 DR Pfam; PF00067; P450; 1.
 DR PRINTS; PR00385; P450; 1.
 DR PROSITE; PS00086; CYTOCHROME P450; 1.
 KW Electron transport; Oxidoreductase; Monooxygenase; Heme;
 Transmembrane.

RA Takano M., Kondo J., Yavama K., Otani M., Sano K., Okamoto H.;
 RT "Molecular cloning of cDNAs for mouse low-molecular-weight and high-
 RL molecular-weight prekallikogens.";
 RN Biochim. Biophys. Acta 1352:222-230 (1997).
 [2]
 RP SEQUENCE FROM N.A. (ISOFORM LMW).
 RC STRAIN=C57BL/6J; TISSUE=Placenta;
 RE MEDLINE=22354683; PubMed=12466851;
 RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
 RA Nikaido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
 RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
 RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
 RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
 RA Blake J.A., Bradt D., Brusci V., Clothia C., Corbani L.E., Cousins S.,
 RA Dalla B., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
 RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
 RA Grimmond S., Gustinchik S., Hirokawa N., Jackson I.J., Jarvis E.D.,
 RA Kanai A., Kawai H., Kawasawa Y., Kedzierski R.M., King B.L.,
 RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
 RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
 RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
 RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
 RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
 RA Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
 RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
 RA Verardo R., Wagner L., Wahlstedt C., Wang Y., Watanabe Y., Wells C.,
 RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
 RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
 RA Hirozane-Kishikawa I., Kono H., Nakamura M., Sakazume N., Sato K.,
 RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
 RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
 RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
 RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
 RA Birney E., Hayashizaki Y.;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs.";
 RL Nature 420:563-573 (2002).
 [3]
 RP SEQUENCE FROM N.A. (ISOFORM LMW).
 RC TISSUE=Liver;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Narisina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Ioshizuka S., Carminci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Paley J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grinstead J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywicki M.I., Skalska U., Smallos D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 CC -!- FUNCTION: (1) Kininogens are inhibitors of thiol proteases, (2)
 CC HMW-kininogen plays an important role in blood coagulation by
 CC helping to position optimally prekallikrein and factor XI next to
 CC factor XII. (3) HMW-kininogen inhibits the thrombin- and plasmin-
 CC induced aggregation of thrombocytes; (4) the active peptide
 CC bradykinin that is released from HMW-kininogen shows a variety of
 CC physiological effects: (4A) influence in smooth muscle
 CC contraction, (4B) induction of hypotension, (4C) natriuresis and
 CC diuresis, (4D) decrease in blood glucose level, (4E) it is a
 CC mediator of inflammation and causes (4E1) increase in vascular
 CC permeability, (4E2) stimulation of nociceptors (4E3) release of
 CC other mediators of inflammation (e.g. prostaglandins), (4F) it has

a cardioprotective effect (directly via bradykinin action,
 indirectly via endothelium-derived relaxing factor action); (5)
 LMW-kininogen inhibits the aggregation of thrombocytes; (6) LMW-
 kininogen is in contrast to HMW-kininogen not involved in blood
 clotting (By similarity).
 CC SUBCELLULAR LOCATION: Secreted.
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event-Alternative splicing; Named isoforms=2;
 CC Name=HMW;
 CC IsoId=008677-1; Sequence=Displayed;
 CC Name=LMW;
 CC IsoId=008677-2; Sequence=VSP_001263, VSP_001264;
 CC TISSUE SPECIFICITY: Plasma.
 CC -!- PTM: Bradykinin is released from kininogen by plasma kallikrein.
 CC -!- SIMILARITY: Contains 3 cystatin-like domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; D84435; BAA19743.1; -;
 CC EMBL; D84415; BAA19742.1; -;
 CC EMBL; AK005547; BAB24115.1; -;
 CC EMBL; BC018158; AAH18158.1; -;
 CC MGD; MGI:1097705; Kng.
 CC InterPro; IPR000010; Cystatin.
 CC InterPro; IPR002395; Kininogen.
 CC Pfam; PF00031; cystatin; 3.
 CC PRINTS; PR00334; KININOGEN.
 CC SMART; SM00043; CY; 3.
 CC PROSITE; P500287; CYSSTATIN; 1.
 CC Glycoprotein; Plasma; Repeat; Thiol protease inhibitor; Vasodilator;
 CC bradykinin; Blood coagulation; Inflammatory response; Signal;
 KW Alternative splicing.
 FT SIGNAL 1 18 POTENTIAL.
 FT CHAIN 19 661 KININOGEN.
 FT CHAIN 19 379 KININOGEN HEAVY CHAIN.
 FT CHAIN 380 388 BRADYKININ.
 FT CHAIN 399 661 KININOGEN LIGHT CHAIN.
 FT DOMAIN 19 135 CYSSTATIN-LIKE 1.
 FT DOMAIN 136 257 CYSSTATIN-LIKE 2.
 FT DOMAIN 258 379 CYSSTATIN-LIKE 3.
 FT DOMAIN 439 524 HIS-RICH.
 FT DISULFID 28 631 INTERCHAIN (BY SIMILARITY).
 FT DISULFID 83 94 BY SIMILARITY.
 FT DISULFID 107 125 BY SIMILARITY.
 FT DISULFID 141 144 BY SIMILARITY.
 FT DISULFID 205 217 BY SIMILARITY.
 FT DISULFID 228 247 BY SIMILARITY.
 FT DISULFID 263 266 BY SIMILARITY.
 FT DISULFID 327 339 BY SIMILARITY.
 FT DISULFID 350 369 BY SIMILARITY.
 FT CARBOHYD 82 82 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 168 168 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 204 204 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 242 242 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT VARSPLIC 401 432 VSPPYIAREQERDAETGQPHGHGHEKQ -> RLRLA
 FT CSVKGLSKAGAPAPQAESSQVKQ (in isoform
 FT LMW).
 FT /FTID=VSP_001263.
 FT Missing (in isoform LMW).
 FT /FTID=VSP_001264.
 FT VARSPLIC 433 661
 FT SEQUENCE 661 AA; 73102 MW; 774460258D58796E CRC64;
 Query Match 43.5%; Score 40.5; DB 1; Length 661;
 Best Local Similarity 40.0%; Pred. No. 64;
 Matches 10; Conservative 1; Mismatches 5; Indels 9; Gaps 1;
 QY 1 GHXKLD-----DDLEHQGH 16

```

Db      471 GHQKLDYLRHQREDGDDHITVGH 495
      ||: ||| ||: ||| ||
RESULT 15
YB53 YEAST
ID YB53 YEAST STANDARD; PRT; 924 AA.
AC P38308;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Hypothetical 104.5 kDa protein in CDC47-KTR3 intergenic region.
GN YBR203W OR YBR1443.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Bacteria; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c;
RX MEDLINE=33377417; PubMed=8368014;
RA Bussereau F., Mallet L., Gaillon L., Jacquet M.;
RT "A 12.8 kb segment, on the right arm of chromosome II from
RT Saccharomyces cerevisiae including part of the DURI2 gene, contains
RT five putative new genes.";
RL Yeast 9:797-806(1993).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z21487; CAA79691.1; -.
DR EMBL; Z36072; CAA85167.1; -.
DR PIR; S34926; S34926.
DR GenOnline; 138746; -.
DR SGD; S0000407; YBR203W.
DR GO; GO:0007165; P:signal transduction; IMP.
DR InterPro; IPR001810; F-box.
DR Pfam; PF00646; F-box; 1.
KW Hypothetical protein.
SQ SEQUENCE 924 AA; 104504 MW; B553A2F32D89BE0D CRC64;
Query Match 43.5%; Score 40.5; DB 1; Length 924;
Best Local Similarity 52.9%; Pred.No.91;
Matches 9; Conservative 1; Mismatches 4; Indels 3; Gaps 1;
QY 3 XXXLDDP---LEHQGH 16
Db 492 KKSDDDFKVLKXDSGH 508

```

Search completed: March 4, 2004, 13:08:07
Job time : 7.75 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2004, 13:02:03 ; Search time 32 Seconds
(without alignments)
167.619 Million cell updates/sec

Title: SEQ-C

Perfect score: 93

Sequence: 1 ghxkxkLDDLEHQGGHV 17

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_25.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phase.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_rvirus.*
16: sp_bacteriap.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Description
1	48	51.6	Q8N2S0 401 4 Q8N2S0
2	48	51.6	Q96PE2 2063 4 Q96PE2
3	47	50.5	Q925A6 541 16 Q925A6
4	47	50.5	Q7UEM6 856 16 Q7UEM6
5	45	48.4	Q94FF6 227 10 Q94FF6
6	45	48.4	Q94FF1 227 10 Q94FF1
7	45	48.4	Q94FE7 227 10 Q94FE7
8	45	48.4	Q94T11 233 10 Q94T11
9	45	48.4	Q81250 236 10 Q81250
10	45	48.4	Q81253 239 10 Q81253
11	45	48.4	Q9SBF1 239 10 Q9SBF1
12	44	47.3	Q87XA8 292 16 Q87XA8
13	44	47.3	Q82FL8 344 16 Q82FL8
14	43	46.2	Q9KTH6 174 16 Q9KTH6
15	43	46.2	Q9S830 459 10 Q9S830
16	43	46.2	Q9X873 489 16 Q9X873

17 43 46.2 540 16 Q82LI0
18 43 46.2 683 9 Q9T215
19 43 46.2 878 16 Q8G4B9
20 43 46.2 1878 10 Q8GIJ7
21 43 46.2 1950 10 Q9AU50
22 42.5 45.7 536 5 Q45994
23 42 45.2 125 16 Q8CYP6
24 42 45.2 156 16 Q87RP9
25 42 45.2 210 16 Q97EL6
26 42 45.2 237 5 Q9XWH4
27 42 45.2 245 5 Q21616
28 42 45.2 289 16 Q88J35
29 42 45.2 483 10 Q8RW10
30 42 45.2 610 16 Q82U49
31 42 45.2 687 5 Q9VXP2
32 42 45.2 718 5 Q9XYI9
33 42 45.2 849 3 Q12205
34 42 45.2 1230 2 P77542
35 42 45.2 1448 2 Q937B6
36 42 45.2 1448 2 Q7X3D4
37 42 45.2 1465 5 Q81223
38 42 45.2 1531 2 Q8RSH6
39 42 45.2 1826 5 Q97255
40 42 45.2 2207 10 Q8LJ53
41 42 45.2 2348 10 Q94D51
42 42 45.2 2747 2 Q9L800
43 42 45.2 3282 16 Q9PEY9
44 42 45.2 3442 16 Q9PB58
45 42 45.2 3455 16 Q9FP06

ALIGNMENTS

RESULT 1

Q8N2S0 ID Q8N2S0 PRELIMINARY; PRT; 401 AA.
AC Q8N2S0;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical protein FLJ90019 (fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_taxID=9606;
RN [1]_taxID=9606;
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RA Isoqai T., Ota T., Nishikawa T., Hayashi K., Otsuki T., Sugiyama T.,
RA Suzuki Y., Nagai K., Sugano S., Ishii S., Kawai-Hio Y., Saito K.,
RA Yamamoto J., Makamatsu A., Nakamura Y., Kojima S., Nagahari K.,
RA Masuho Y., Ono T., Okano K., Yoshikawa Y., Aotsuka S., Sasaki N.,
RA Hattori A., Okumura K., Iwayanagi T., Ninomiya K.;
RT "NEDO human cDNA sequencing project."
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK074500; BAC11025.1; -;
KW Hypothetical protein.
FT NON_TER 401 401
SQ SEQUENCE 401 AA; 42191 MW; 9B64941BC7942433 CRC64;

Query Match 51.6%; Score 48; DB 4; Length 401;
Best Local Similarity 50.0%; Pred. No. 11;
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 1 GHXXKLDLLEHQGGH 16

Db 272 GHSSGSDDDRDGGGH 287

||||| : |||

RESULT 2

Q96PE2 ID Q96PE2 PRELIMINARY; PRT; 2063 AA.

```

AC Q96PE2;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Tumor endothelial marker 4.
GN TEM4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21443268; PubMed=11559528;
RA Carson-Walter E.B., Watkins D.N., Nanda A., Vogelstein B.,
RA Kinzler K.W., St Croix B.
RT "Cell surface tumor endothelial markers are conserved in mice and
RT humans.";
RL Cancer Res. 61:6649-6655(2001).
RL EMBL; AF378754; AAL11991.1; -.
DR InterPro; IPR000219; RhoGEF.
DR Pfam; PF00621; RhoGEF; 1.
DR SMART; SM00325; RhoGEF; 1.
DR PROSITE; PS00010; DH 2; 1.
SQ SEQUENCE 2063 AA; 221671 MW; 76A53FF6CBFF5C0 CRC64;

Query Match 51.6%; Score 48; DB 4; Length 2063;
Best Local Similarity 50.0%; Pred. No. 66;
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 GHXKLDLDDLEHGGHV 16
DB 272 GHSSGSDDRDGGGH 287

RESULT 3
ID Q9Z5A6 PRELIMINARY; PRT; 541 AA.
AC Q9Z5A6;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Putative long-chain-fatty-acid-CoA ligase.
GN SC06196 OR SC2G5.17.
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomyces; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Oliver K., Harris D.;
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Bentley S.D., Parkhill J., Barrell B.G., Rajandream M.A.;
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RX MEDLINE=97000351; PubMed=8843436;
RA Redenbach M., Kieser H.M., Denapante D., Eichner A., Cullum J.,
RA Kinashi H., Hopwood D.A.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL Mol. Microbiol. 21:77-96(1996).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RX MEDLINE=21996410; PubMed=12000953;
RA Bentley S.D., Chater K.F., Cerdano-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,

```

```

RA Cronin A., Praeger A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neil S.,
RA Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2).";
RL Nature 417:141-147(2002).
DR EMBL; AL939126; CAB36604.1; -.
DR PIR; T34850; T34850.
DR DR HSP; P08659; ILCI.
DR GO; GO:0016874; F.ligase activity; IEA.
DR GO; GO:000152; P.metabolism; IEA.
DR InterPro; IPR000873; AMP-bind.
DR Pfam; PF00501; AMP-binding; 1.
DR PROSITE; PS00455; AMP_BINDING; 1.
KW Ligase; Complete proteome.
SQ SEQUENCE 541 AA; 59034 MW; CEB7374431F28CE5 CRC64;

Query Match 50.5%; Score 47; DB 16; Length 541;
Best Local Similarity 53.3%; Pred. No. 22;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 3 KXKLDLDDLEHGGHV 17
DB 346 QTRMDDDLHRTGTG 360

RESULT 4
ID Q7UEM6 PRELIMINARY; PRT; 856 AA.
AC Q7UEM6;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Serine/threonine protein kinase (EC 2.7.1.1.-).
GN RB11244.
OS Rhodopirellula baltica.
OC Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales;
OC Planctomycetaceae; Pirellula.
OX NCBI_TaxID=117;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1;
RX MEDLINE=22735913; PubMed=12835416;
RA Gloeckner P.O., Kube M., Bauer M., Teeling H., Lombardot T.,
RA Ludwig W., Gade D., Beck A., Borzym K., Heitmann K., Rabus R.,
RA Schlesner H., Amann R., Reinhardt R.;
RT "Complete genome sequence of the marine planctomycete Pirellula sp.
RT strain 1.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:8298-8303(2003).
DR EMBL; BX294153; CAD73009.1; -.
KW Kinase; Serine/threonine-protein kinase; Transferase;
KW Complete proteome.
SQ SEQUENCE 856 AA; 95464 MW; A71CAF206878C42F CRC64;

Query Match 50.5%; Score 47; DB 16; Length 856;
Best Local Similarity 50.0%; Pred. No. 36;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 GHXKLDLDDLEHGG 14
DB 675 GHRVTLDDSIKHS 688

RESULT 5
ID Q94FF6 PRELIMINARY; PRT; 227 AA.
AC Q94FF6;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

```


DE Globulin 1 (Fragment).
 OS Zea mays subsp. mays (maize).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC PACCAD clade; Panicoideae; Andropogoneae; Zea.
 OX NCBI_TaxID=4578;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Guirua;
 RA Tenaillon M.I., Sawkins M.C., Gaut R.L., Long A.D., Doebley J.F.,
 RA Gaut B.S.;
 RT "Patterns of DNA sequence diversity along chromosome 1 of maize (Zea
 mays subsp. mays L.).";
 RL Proc. Natl. Acad. Sci. U.S.A. 96:9161-9166 (2001).
 DR EMBL; AF377673; AAK60212.1; -.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0045735; F:nutrient reservoir activity; IEA.
 DR InterPro; IPR005479; CPase_L_D2.
 DR InterPro; IPR006045; Cupin.
 DR Pfam; PF00190; Cupin; 1.
 DR PROSITE; PS00867; CPSASE_2; 1.
 FT NON TER 227 227
 SQ SEQUENCE 227 AA; 26139 MW; 5A23112B31C7F645 CRC64;

Query Match 48.4%; Score 45; DB 10; Length 227;
 Best Local Similarity 70.0%; Pred. No. 18;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 7 DDDLEHQGGH 16
 ||:|:|
 Db 26 DDNLHHGGH 35

RESULT 6

Q94FF1 PRELIMINARY; PRT; 227 AA.
 AC Q94FF1;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Globulin 1 (Fragment).
 OS Zea mays subsp. mays (maize).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC PACCAD clade; Panicoideae; Andropogoneae; Zea.
 OX NCBI_TaxID=4578;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Sabanero;
 RA Tenaillon M.I., Sawkins M.C., Gaut R.L., Long A.D., Doebley J.F.,
 RA Gaut B.S.;
 RT "Patterns of DNA sequence diversity along chromosome 1 of maize (Zea
 mays subsp. mays L.).";
 RL Proc. Natl. Acad. Sci. U.S.A. 96:9161-9166 (2001).
 DR EMBL; AF377673; AAK60217.1; -.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0045735; F:nutrient reservoir activity; IEA.
 DR InterPro; IPR005479; CPase_L_D2.
 DR InterPro; IPR006045; Cupin.
 DR Pfam; PF00190; Cupin; 1.
 DR PROSITE; PS00867; CPSASE_2; 1.
 FT NON TER 227 227
 SQ SEQUENCE 227 AA; 26097 MW; B465266E142F8CE3 CRC64;

Query Match 48.4%; Score 45; DB 10; Length 227;
 Best Local Similarity 70.0%; Pred. No. 18;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 7 DDDLEHQGGH 16
 ||:|:|
 Db 26 DDNLHHGGH 35

RESULT 7

Q94FE7 PRELIMINARY; PRT; 227 AA.
 AC Q94FE7;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Globulin 1 (Fragment).
 OS Zea mays subsp. mays (maize).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC PACCAD clade; Panicoideae; Andropogoneae; Zea.
 OX NCBI_TaxID=4578;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Texas 601;
 RA Tenaillon M.I., Sawkins M.C., Gaut R.L., Long A.D., Doebley J.F.,
 RA Gaut B.S.;
 RT "Patterns of DNA sequence diversity along chromosome 1 of maize (Zea
 mays subsp. mays L.).";
 RL Proc. Natl. Acad. Sci. U.S.A. 96:9161-9166 (2001).
 DR EMBL; AF377689; AAK60227.1; -.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0045735; F:nutrient reservoir activity; IEA.
 DR InterPro; IPR005479; CPase_L_D2.
 DR InterPro; IPR006045; Cupin.
 DR Pfam; PF00190; Cupin; 1.
 DR PROSITE; PS00867; CPSASE_2; 1.
 FT NON TER 227 227
 SQ SEQUENCE 227 AA; 26113 MW; B465266E0787B673 CRC64;

Query Match 48.4%; Score 45; DB 10; Length 227;
 Best Local Similarity 70.0%; Pred. No. 18;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 7 DDDLEHQGGH 16
 ||:|:|
 Db 26 DDNLHHGGH 35

RESULT 8

Q9AT11 PRELIMINARY; PRT; 233 AA.
 AC Q9AT11;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Globulin 1 (Fragment).
 GN GLBL.
 OS Zea diploperennis (Diploperennial teosinte).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC PACCAD clade; Panicoideae; Andropogoneae; Zea.
 OX NCBI_TaxID=4576;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=dlia, and cv. M005;
 RX MEDLINE=21231628; PubMed=11333248;
 RA Tiffin P., Gaut B.S.;
 RT "sequence Diversity in the Tetraploid Zea perennis and the Closely
 Related Diploid Z. diploperennis. Insights from four nuclear loci.";
 RL Genetics 158:401-412 (2001).
 DR EMBL; AF329790; AAK20332.1; -.
 DR HSP; P50477; 1CAU.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0045735; F:nutrient reservoir activity; IEA.
 DR InterPro; IPR005479; CPase_L_D2.
 DR InterPro; IPR006045; Cupin.
 DR Pfam; PF00190; Cupin; 1.
 DR PROSITE; PS00867; CPSASE_2; 1.
 FT NON TER 233 233
 SQ SEQUENCE 233 AA; 26794 MW; 5B9553BA3935CB75 CRC64;

Query Match 48.4%; Score 45; DB 10; Length 233;
 Best Local Similarity 70.0%; Pred. No. 19;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 7 DDDLEHQGGH 16
 ||:|||||
 DB 26 DDNLHHGGH 35

RESULT 9

ID O81250 PRELIMINARY; PRT; 236 AA.
 AC O81250;
 DT 01-NOV-1998 (TREMBLrel. 08, Created)
 DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
 DE Globulin-1 (Fragment).
 OS Zea mays subsp. mays (maize).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC PACAD clade; Panicoideae; Andropogoneae; Zea.
 CX NCBI_TaxID=4578;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hilton H., Gaut B.S.;
 RT "Speciation and domestication in maize and its wild relatives:
 evidence from the Globulin-1 gene.";
 RL Genetics 0:0-0(1998).
 DR EMBL; AF064213; AAC31456.1; -.
 DR HSSP; P50477; 1CAU.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0045735; F:nutrient reservoir activity; IEA.
 DR InterPro; IPR005479; CPase_L_D2.
 DR InterPro; IPR006045; Cupin_L.
 DR Pfam; PF00190; Cupin; 1.
 DR PROSITE; PS00867; CPASE_2; 1.
 FT NON TER 236
 FT SEQUENCE 236 AA; 27019 MW; 1F3D9BD92C032E05 CRC64;

Query Match 48.4%; Score 45; DB 10; Length 236;
 Best Local Similarity 70.0%; Pred. No. 19;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 7 DDDLEHQGGH 16
 ||:|||||
 DB 26 DDNLHHGGH 35

RESULT 10

ID O81253 PRELIMINARY; PRT; 239 AA.
 AC O81253;
 DT 01-NOV-1998 (TREMBLrel. 08, Created)
 DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
 DE Globulin-1 (Fragment).
 OS Zea mays subsp. mays (maize).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC PACAD clade; Panicoideae; Andropogoneae; Zea.
 CX NCBI_TaxID=4578;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hilton H., Gaut B.S.;
 RT "Speciation and domestication in maize and its wild relatives:
 evidence from the Globulin-1 gene.";
 RL Genetics 0:0-0(1998).
 DR EMBL; AF064216; AAC31459.1; -.
 DR HSSP; P50477; 1CAU.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0045735; F:nutrient reservoir activity; IEA.
 DR InterPro; IPR005479; CPase_L_D2.
 DR InterPro; IPR006045; Cupin.

DR Pfam; PF00190; Cupin; 1.
 DR PROSITE; PS00867; CPASE_2; 1.
 FT NON TER 239
 FT SEQUENCE 239 AA; 27384 MW; 628924A8D7BA7773 CRC64;

Query Match 48.4%; Score 45; DB 10; Length 239;
 Best Local Similarity 70.0%; Pred. No. 19;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 7 DDDLEHQGGH 16
 ||:|||||
 DB 26 DDNLHHGGH 35

RESULT 11

ID O9S8F1 PRELIMINARY; PRT; 239 AA.
 AC O9S8F1;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
 DE Globulin-1 (Fragment).
 OS Zea mays subsp. mays (maize).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC PACAD clade; Panicoideae; Andropogoneae; Zea.
 CX NCBI_TaxID=4578;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hilton H., Gaut B.S.;
 RT "Speciation and domestication in maize and its wild relatives:
 evidence from the Globulin-1 gene.";
 RL Genetics 0:0-0(1998).
 DR EMBL; AF064218; AAC31461.1; -.
 DR HSSP; P50477; 1CAU.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0045735; F:nutrient reservoir activity; IEA.
 DR InterPro; IPR005479; CPase_L_D2.
 DR InterPro; IPR006045; Cupin_L.
 DR Pfam; PF00190; Cupin; 1.
 DR PROSITE; PS00867; CPASE_2; 1.
 FT NON TER 239
 FT SEQUENCE 239 AA; 27499 MW; 147C4F61F65307FA CRC64;

Query Match 48.4%; Score 45; DB 10; Length 239;
 Best Local Similarity 70.0%; Pred. No. 19;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 7 DDDLEHQGGH 16
 ||:|||||
 DB 26 DDNLHHGGH 35

RESULT 12

ID O87XA8 PRELIMINARY; PRT; 292 AA.
 AC O87XA8;
 DT 01-JUN-2003 (TREMBLrel. 24, Created)
 DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
 DE Membrane protein, putative.
 GN PSPT04280.
 OS Pseudomonas syringae (pv. tomato).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Pseudomonas.
 CX NCBI_TaxID=323;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=DC3000;
 RA Buehl R., Joardar V., Khouri H., Fedorova N., Tran B., Russell D.,
 RA Berry K., Utterback T., Van Aken S., Feldblyum T., Gwinn M.,
 RA Dodson R., DeBoy R., Durkin A., Kolonay J., Madupu R., Daugherty S.,
 RA Brinkac L., Beanan M., Haft D., Selengut J., Nelson W., Davidsen T.,

```

RA White O., Fraser C., Collmer A.;
RT "Complete sequence of Pseudomonas syringae.";
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE016871; AAC57731.1; -.
DR TIGR; PSP04280; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0015099; F:nickel ion transporter activity; IEA.
DR GO; GO:0015675; P:nickel ion transport; IEA.
DR InterPro; IPR004688; NiCo.
DR Pfam; PF03824; NiCo; 1.
DR Complete proteome.
SQ SEQUENCE 252 AA; 31560 MW; B456C00B314B5ED7 CRC64;

Query Match 47.3%; Score 44; DB 16; Length 292;
Best Local Similarity 53.3%; Pred. No. 35;
Matches 8; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 2 HXKXLDLHGGGH 16
DB 136 HSHNHDLHGGHH 150

RESULT 13
Q82FL8 PRELIMINARY; PRT; 344 AA.
AC Q82FL8;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative dioxygenase.
GN SAV4234.
OS Streptomyces avermitilis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomycetes.
OX NCBI_TaxID=33903;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=21477403; PubMed=11572946;
RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osonoe T.,
RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
RT "Genome sequence of an industrial microorganism Streptomyces
RT avermitilis: deducing the ability of producing secondary
RT metabolites.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
RV [2]
RP SEQUENCE FROM N.A.
RC STRAIN=WA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=22608306; PubMed=12692562;
RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
RA Sakaki Y., Hattori M., Omura S.;
RT "Complete genome sequence and comparative analysis of the industrial
RT microorganism Streptomyces avermitilis.";
RL Nat. Biotechnol. 21:526-531(2003).
RR EMBL; AF005038; BAC71946.1; -.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0008199; F:ferric iron binding; IEA.
DR GO; GO:0016702; F:oxidoreductase activity, acting on single d. . .; IEA.
DR GO; GO:0006520; P:amino acid metabolism; IEA.
DR GO; GO:0006725; P:aromatic compound metabolism; IEA.
DR InterPro; IPR006034; Asp/Glutamase.
DR InterPro; IPR000627; Dioxygenase.
DR InterPro; IPR006311; Tat.
DR Pfam; PF00775; Dioxygenase; 1.
DR TIGRfam; TIGR01409; TAT signal seq; 1.
DR PROSITE; PS00144; ASN_GLM ASE_1; 1.
KW Dioxygenase; Complete proteome.
SQ SEQUENCE 344 AA; 35028 MW; 5878612258C334CF CRC64;

Query Match 47.3%; Score 44; DB 16; Length 344;
Best Local Similarity 53.3%; Pred. No. 42;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

RA White O., Fraser C., Collmer A.;
RT "Complete sequence of Pseudomonas syringae.";
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE016871; AAC57731.1; -.
DR TIGR; PSP04280; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0015099; F:nickel ion transporter activity; IEA.
DR GO; GO:0015675; P:nickel ion transport; IEA.
DR InterPro; IPR004688; NiCo.
DR Pfam; PF03824; NiCo; 1.
DR Complete proteome.
SQ SEQUENCE 252 AA; 31560 MW; B456C00B314B5ED7 CRC64;

Query Match 46.2%; Score 43; DB 16; Length 174;
Best Local Similarity 58.3%; Pred. No. 29;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 6 LDDLEHGGGHV 17
DB 88 LNDWLHGGHHI 99

RESULT 15
Q8SS30 PRELIMINARY; PRT; 459 AA.
AC Q8SS30;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative kinesin-like centromere protein.
GN F14P13.22.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B.,
RA Rensing C.M., Koo H., Fujii C.Y., Utterback T.R., Barnstead M.E.,
RA Bowman C.L., White O., Nierman W.C., Fraser C.M.;
RT "Arabidopsis thaliana chromosome III BAC F14P13 genomic sequence.";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC009400; AAF02823.1; -.
DR HSP; P33176; 1BQ2.
DR GO; GO:0005871; C:kinesin complex; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0003774; F:motor activity; IEA.
DR GO; GO:0007017; P:microtubule-based process; IEA.

```

```

DR InterPro: IPR001752; kinesin_motor.
DR Pfam: P20225; kinesin; 1.
DR PRINTS: PR00380; KINESINHEAVY.
DR SMART: SM00129; KISC; 1.
DR PROSITE: PS00411; KINESIN MOTOR_DOMAIN1; 1.
DR PROSITE: PS00067; KINESIN MOTOR_DOMAIN2; 1.
SQ SEQUENCE 459 AA; 51872 MW; 7C8487E9B7038E6A CRC64;

Query Match          46.2%; Score 43; DB 10; Length 459;
Best Local Similarity 61.5%; Pred. No. 87;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 5 KLDDDLHQGGHV 17
   || : |||||
Db 267 KLSEGVETQGGHV 279

```

Search completed: March 4, 2004, 13:10:29
Job time : 34 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2004, 12:52:17 ; Search time 16.6765 Seconds
(without alignments)
101.657 Million cell updates/sec

Title: SEQ-D
Perfect score: 29
Sequence: 1 ghkxkl 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A: Geneseqp29Jan04: *
1: Geneseqp1980s: *
2: Geneseqp1990s: *
3: Geneseqp2000s: *
4: Geneseqp2001s: *
5: Geneseqp2002s: *
6: Geneseqp2003as: *
7: Geneseqp2003bs: *
8: Geneseqp2004s: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	28	96.6	218	4	Aae02538 A. thalia
2	28	96.6	218	6	Ada15595 A. thalia
3	28	96.6	218	7	Add30748 Plant yie
4	28	96.6	247	3	Aag29968 Arabidops
5	28	96.6	304	3	Aag29967 Arabidops
6	28	96.6	304	6	Ada15475 A. thalia
7	28	96.6	407	6	Abu08934 Human tum
8	27	93.1	13	6	Aao30886 Peptide m
9	27	93.1	16	3	Aay81998 Human two
10	27	93.1	16	5	Abb78700 Human kin
11	27	93.1	28	3	Aay81996 Human hig
12	27	93.1	47	3	Aay913345 Light cha
13	27	93.1	55	3	Aay93346 Light cha
14	27	93.1	62	3	Aay93348 Light cha
15	27	93.1	63	2	Aar75186 Partial p
16	27	93.1	83	3	Aay93347 Light cha
17	27	93.1	94	3	Aay93351 Light cha
18	27	93.1	100	4	Abg29190 Novel hum
19	27	93.1	125	5	Abb78708 Human hig
20	27	93.1	128	4	Aag74402 Human col
21	27	93.1	131	2	Aar75181 Partial p
22	27	93.1	139	4	Aao01620 Human pol
23	27	93.1	158	5	Abb78709 Calmoduli
24	27	93.1	170	7	Adb80073 Mycobacte
25	27	93.1	177	3	Aab33089 Pinus rad

ALIGNMENTS

RESULT 1		
AAE02538		
ID	AAE02538	standard; protein; 218 AA.
XX	AC	AAE02538;
XX	AC	AAE02538;
DT	10-AUG-2001	(first entry)
DE	A. thaliana transcription factor G36.	
XX	Plant transcription factor; phenotype; sugar sensing characteristic;	
KW	transgenic plant; plant yield; growth; germination; photosynthesis;	
KW	glyoxylate metabolism; respiration; pathogen response; wounding response;	
KW	cell cycle regulation; pigmentation; flowering; senescence; physiology;	
KW	storage organ; metabolism.	
XX	Arabidopsis thaliana.	
OS	XX	
FH	Key	Location/Qualifiers
FT	Domain	67..134
FT	/note= "Conserved domain"	
XX	WO200135725-A1.	
XX	25-MAY-2001.	
PF	14-NOV-2000; 2000WO-US031414.	
XX	17-NOV-1999; 99US-0166228P.	
PR	17-APR-2000; 2000US-0197899P.	
PR	22-AUG-2000; 2000US-0227439P.	
XX	(MEND-) MENDEL BIOTECHNOLOGY INC.	
PA	(JIAN/) JIANG C.	
PA	(HEAR/) HEARD J.	
PA	(PINE/) PINEDA O.	
PA	(PILG/) PILGRIM M.	
PA	(ADAM/) ADAM L.	
PA	(RIEC/) RIECHMANN J L.	
PA	(YUGG/) YU G.	
PA	(SAMA/) SAMAHA R.	
XX	Jiang C, Heard J, Pineda O, Pilgrim M, Adam L, Riechmann JL;	
PI	Yu G, Samaha R;	
XX	WPI; 2001-335977/35.	
DR	N-PSDB; AAD06639.	
XX		

Aay93353 Light cha
Aay93349 Light cha
Abg08012 Novel hum
Aag27712 Arabidops
Aay93342 Light cha
Aab73620 Human hig
Abg21100 Novel hum
Aag53316 Arabidops
Abr41202 Human DIT
Abd46064 rapE sequ
Abu99146 Novel hum
Aay19819 B. burgdo
Aag27711 Arabidops
Aay19818 B. burgdo
Abu19366 Protein e
Aae09718 Human ubi
Aar06519 Microspor
Aaw08378 Brassica
Aag53315 Arabidops
Abb92296 Herbicida

26 27 93.1 179 3 AAY93353
27 27 93.1 186 3 AAY93349
28 27 93.1 240 4 ABG08012
29 27 93.1 243 3 AAG27712
30 27 93.1 255 3 AAY93342
31 27 93.1 255 4 AAB73620
32 27 93.1 305 4 ABG21100
33 27 93.1 342 3 AAG53316
34 27 93.1 357 7 ABR41202
35 27 93.1 375 6 ABD46064
36 27 93.1 415 6 ABU99146
37 27 93.1 446 2 AAY19819
38 27 93.1 452 3 AAG27711
39 27 93.1 482 2 AAY19818
40 27 93.1 482 6 ABU19366
41 27 93.1 485 4 AAE09718
42 27 93.1 509 2 AAR06519
43 27 93.1 509 2 AAW08378
44 27 93.1 551 3 AAG53315
45 27 93.1 551 5 ABB92296

PT Nucleic acids encoding plant transcription factor polypeptides, useful
PT for altering the sugar sensing characteristics of plants and increasing
PT yield, e.g. corn, potato and cotton plants.
XX Claim 4; Page 57-58; 151pp; English.
XX The patent relates to polynucleotides encoding 35 plant transcription
CC factors which may be used to modify phenotype associated with a plant's
CC sugar sensing characteristics and increasing yield when their expression
CC level is altered. Sugars are central regulatory molecules that control
CC aspects of physiology, metabolism and development. Therefore the cDNAs
CC and proteins of the invention are useful for modifying the growth and
CC germination rates of plants, photosynthesis, glyoxylate metabolism,
CC respiration, starch and sucrose synthesis and degradation, pathogen
CC response, wounding response, cell cycle regulation, pigmentation,
CC flowering and senescence of plants and for modifying sink-source,
CC relationships in seeds, tubers, roots, and other storage organs leading
CC to an increase in yield. The transcription factor polynucleotides and
CC polypeptides may be used to alter the structure and developmental
CC characteristics of plants such as soybean, wheat, corn, potato, cotton,
CC rice, oilseed, sunflower, alfalfa, sugarcane, turf, banana, blackberry,
CC blueberry, strawberry, raspberry, cantaloupe, carrot, cauliflower,
CC coffee, cucumber, eggplant, grapes, honey dew, lettuce, mango, melon,
CC onion, papaya, peas, peppers, pineapple, spinach, squash, sweet corn,
CC tobacco, tomato, watermelon, roseaceous fruits and/or vegetable brassicas.
CC The present sequence is an Arabidopsis thaliana transcription factor
XX
SQ Sequence 218 AA;

Query Match 96.6%; Score 28; DB 4; Length 218;
Best Local Similarity 83.3%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GHKXKL 6
||| ||
Db 118 GHKAKL 123

RESULT 2

ADAI15595
ID ADAI15595 standard; protein; 218 AA.

XX AC ADAI15595;

XX DT 06-NOV-2003 (first entry)

XX DE A. thaliana protein, SEQ ID NO:150.

XX KW Transgenic plant; plant; transcription factor; trait; root;
XX plant stability.

XX OS Arabidopsis thaliana.

XX PN US2003061637-A1.

XX PD 27-MAR-2003.

XX PF 21-OCT-2002; 2002US-00278173.

XX PR 23-MAR-1999; 99US-0125814P.

XX PR 22-MAR-2000; 2000US-00533393.

XX PA (JIAN/) JIANG C.
XX PA (BROU/) BROUN P.
XX PA (RIEC/) RIECHMANN J L.
XX PA (PINE/) PINEDA O.
XX PA (ZHAN/) ZHANG J.
XX PA (YUGG/) YU G.
XX PA (PILG/) PILGRIM M.
XX PA (KEDD/) KEDDIE J.
XX PA (HEAR/) HEARD J.
XX PA (REUB/) REUBER L.
XX PA (RATC/) RATCLIFFE O.

PA (ADAM/) ADAM L.
PA (SAMA/) SAMAHA R.
XX JIANG C, BROUN P, RIECHMANN JL, PINEDA O, ZHANG J, YU G,
XX PILGRIM M, KEDDIE J, HEARD J, REUBER L, RATCLIFFE O, ADAM L;
XX Samaha R;
XX WPI: 2003-555503/52.
XX N-PSDB; ADAI15594.
XX New transgenic plant comprising a recombinant polynucleotide, useful for
XX altering a plant's trait for increasing plant stability.
XX Claim 1; SEQ ID NO 150; 159pp; English.

XX The invention discloses a transgenic plant having an A. thaliana
CC recombinant polynucleotide (a transcription factor) comprising a sequence
CC encoding a polypeptide with at least 6 consecutive amino acids of one of
CC the amino acid sequences given in the specification. The recombinant
CC polynucleotide alters a trait of the transgenic plant's roots when
CC compared to the same trait of the roots of another plant lacking the
CC recombinant polynucleotide. Also claimed are methods for altering the
CC expression levels of at least one gene of a plant, altering a trait
CC associated with a plant's roots and altering a plant's trait. The
CC recombinant polynucleotide further comprises a promoter operably linked
CC to the nucleotide sequence. The promoter is constitutive or inducible or
CC root-active. The method for altering a trait associated with roots
CC comprises transforming a plant with the recombinant polynucleotide,
CC selecting the transformed plants and identifying a transformed plant with
CC roots having an altered trait. The method for altering the expression
CC levels of at least one gene of a plant comprises transforming a plant
CC with the recombinant polynucleotide and selecting the transformed plant.
CC The method for altering a trait associated with a plant's roots comprises
CC transforming a plant with the recombinant polynucleotide and selecting
CC the transformed plant. The method for altering a plant's trait comprises
CC providing a database sequence comparing the database sequence with the
CC polypeptide, selecting the database that meets selected sequence criteria
CC and transforming the selected database sequence in the plant. The methods
CC also comprise providing a test polynucleotide, hybridising the test
CC polynucleotide with the recombinant polynucleotide at low stringency and
CC transforming the hybridising test polynucleotide in a plant to alter a
CC trait of the plant. The transgenic plant is useful for altering a plant's
CC trait for increasing plant stability. The sequence presented is a protein
CC of the invention.

XX SQ Sequence 218 AA;

Query Match 96.6%; Score 28; DB 6; Length 218;
Best Local Similarity 83.3%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GHKXKL 6
||| ||
Db 118 GHKAKL 123

RESULT 3

ADD30748

ID ADD30748 standard; protein; 218 AA.

XX AC ADD30748;

XX DT 15-JAN-2004 (first entry)

XX DE Plant yield-related protein from clone G26.

XX KW transcription factor; transgenic plant; growth rate; senescence;
XX seed germination rate; plant vigor; seedling vigor.

XX OS Arabidopsis thaliana.

XX PN WO2003013227-A2.

```

PD 20-FEB-2003.
PF 09-AUG-2002; 2002WO-USO25805.
XX
PR 09-AUG-2001; 2001US-0310847P.
PR 19-NOV-2001; 2001US-0336049P.
PR 11-DEC-2001; 2001US-0338692P.
PR 14-JUN-2002; 2002US-00171469.
XX
PA (MEND-) MENDEL BIOTECHNOLOGY INC.
XX
PI Ratcliffe O, Riechmann JL, Adam LJ, Dubell AT, Heard JB;
PI Pilgrim ML, Jiang C, Reuber TL, Creelman RA, Pineda O, Yu G;
PI Brown PE;
XX
DR WPI; 2003-248221/24.
DR N-PSDB; ADD30747.
XX
XX New plant transcription factor polynucleotides and polypeptides, useful
PT in producing transgenic plants with commercially valuable properties,
PT such as an alteration in a plant growth characteristic, e.g. growth rate
PT or apomixis.
XX
PS Disclosure; SEQ ID NO 777; 454pp; English.
XX
CC The invention relates to a number of isolated Arabidopsis thaliana cDNA
CC sequences and their encoded proteins which are especially transcription
CC factor related cDNA's and proteins. The isolated or recombinant plant
CC transcription factor polynucleotides and polypeptides are useful in
CC producing transgenic plants with commercially valuable properties, i.e.
CC modified or altered desirable traits as compared to a reference plant,
CC such as an alteration in a plant growth characteristic, e.g. growth rate,
CC germination rate of seeds, vigor of plants and seedlings, or leaf and
CC flower senescence. Sequence information related to the polynucleotides
CC and polypeptides can also be used in bioinformatic search methods. The
CC transgenic plant is useful for growing a progeny plant from a parent
CC plant. This sequence represents one of the proteins of the invention.
XX
SQ Sequence 218 AA;
    Query Match          96.6%; Score 28; DB 7; Length 218;
    Best Local Similarity 83.3%; Pred.No. 1.2e-02;
    Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY      1 GHKXKL 6
Db      |||||
      118 GHKAKL 123
RESULT 4
AAG29968
ID AAG29968 standard; protein; 247 AA.
XX
AC AAG29968;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 35744.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
XX 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR

```

```

PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135829P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR

```

```

PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149375P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 23-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 25-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0156659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.

PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 18-OCT-1999; 99US-0159638P.
PR 21-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160800P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 96.6%; Score 28; DB 3; Length 247;
Best Local Similarity 83.3%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GHXXKL 6
Db 149 GHKAKL 154

RESULT 5
AAG29967
ID AAG29967 standard; protein; 304 AA.
XX
AC AAG29967;
XX
XX 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 35743.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.

```


PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-01334256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139452P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 02-JUL-1999; 99US-0142154P.
PR 06-JUL-1999; 99US-0142055P.
PR 08-JUL-1999; 99US-0142390P.
PR 09-JUL-1999; 99US-0142803P.
PR 12-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144844P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145132P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 18-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-015753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.

PR 14-OCT-1999; 99US-0159329P.
 PR 14-OCT-1999; 99US-0159330P.
 PR 14-OCT-1999; 99US-0159331P.
 PR 14-OCT-1999; 99US-0159637P.
 PR 14-OCT-1999; 99US-0159638P.
 PR 18-OCT-1999; 99US-0159584P.
 PR 21-OCT-1999; 99US-0160741P.
 PR 21-OCT-1999; 99US-0160767P.
 PR 21-OCT-1999; 99US-0160768P.
 PR 21-OCT-1999; 99US-0160770P.
 PR 21-OCT-1999; 99US-0160814P.
 PR 21-OCT-1999; 99US-0160815P.
 PR 22-OCT-1999; 99US-0160980P.
 PR 22-OCT-1999; 99US-0160981P.
 PR 22-OCT-1999; 99US-0160989P.
 PR 25-OCT-1999; 99US-0161404P.
 PR 25-OCT-1999; 99US-0161405P.
 PR 25-OCT-1999; 99US-0161406P.
 PR 26-OCT-1999; 99US-0161359P.
 PR 26-OCT-1999; 99US-0161360P.
 PR 28-OCT-1999; 99US-0161920P.
 PR 28-OCT-1999; 99US-0161923P.
 PR 28-OCT-1999; 99US-0161931P.
 PR 29-OCT-1999; 99US-0162142P.

Query Match 96.6%; Score 28; DB 3; Length 304;
 Best Local Similarity 83.3%; Pred. No. 1.7e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GHXXKL 6
 DB 206 GHKAKL 211

RESULT 6
 ID ADA15475 standard; protein; 304 AA.
 XX ADA15475;
 XX 06-NOV-2003 (first entry)
 DT A. thaliana protein, SEQ ID NO:30.
 DE Transgenic plant; plant; transcription factor; trait; root;
 XX plant stability.
 XX Arabidopsis thaliana.
 XX US2003061637-A1.
 XX 27-MAR-2003.
 XX 21-OCT-2002; 2002US-00278173.
 XX 23-MAR-1999; 99US-0125814P.
 XX 22-MAR-2000; 2000US-00533393.
 XX (JIAN//) JIANG C.
 PA (BROU//) BROUN P.
 PA (RIEC//) RIECHMANN J L.
 PA (PINE//) PINEDA O.
 PA (ZHAN//) ZHANG J.
 PA (YUGG//) YU G.
 PA (PILG//) PILGRIM M.
 PA (KEDD//) KEDDIE J.
 PA (HEAR//) HEARD J.
 PA (REUB//) REUBER L.
 PA (RATC//) RATCLIFFE O.
 PA (ADAM//) ADAM L.
 PA (SAMA//) SAMAHA R.
 XX

PI Jiang C, Broun P, Riechmann JL, Pineda O, Zhang J, Yu G,
 PI Pilgrim M, Keddie J, Heard J, Reuber L, Ratcliffe O, Adam L,
 XX Samaha R;
 DR WPI; 2003-555503/52.
 DR N-FSDB; ADA15474.
 XX
 PT New transgenic plant comprising a recombinant polynucleotide, useful for
 PT altering a plant's trait for increasing plant stability.
 XX
 XX Claim 1; SEQ ID NO 30; 159pp; English.
 XX
 CC The invention discloses a transgenic plant having an A. thaliana
 CC recombinant polynucleotide (a transcription factor) comprising a sequence
 CC encoding a polypeptide with at least 6 consecutive amino acids of one of
 CC the amino acid sequences given in the specification. The recombinant
 CC polynucleotide alters a trait of the transgenic plant's roots when
 CC compared to the same trait of the roots of another plant lacking the
 CC recombinant polynucleotide. Also claimed are methods for altering the
 CC expression levels of at least one gene of a plant, altering a trait
 CC associated with a plant's roots and altering a plant's trait. The
 CC recombinant polynucleotide further comprises a promoter operably linked
 CC to the nucleotide sequence. The promoter is constitutive or inducible or
 CC root-active. The method for altering a trait associated with roots
 CC comprises transforming a plant with the recombinant polynucleotide,
 CC selecting the transformed plants and identifying a transformed plant with
 CC roots having an altered trait. The method for altering the expression
 CC levels of at least one gene of a plant comprises transforming a plant
 CC with the recombinant polynucleotide and selecting the transformed plant.
 CC The method for altering a trait associated with a plant's roots comprises
 CC transforming a plant with the recombinant polynucleotide and selecting
 CC the transformed plant. The method for altering a plant's trait comprises
 CC providing a database sequence comparing the database sequence with the
 CC polypeptide, selecting the database that meets selected sequence criteria
 CC and transforming the selected database sequence in the plant. The methods
 CC also comprise providing a test polynucleotide, hybridising the test
 CC polynucleotide with the recombinant polynucleotide at low stringency and
 CC transforming the hybridising test polynucleotide in a plant to alter a
 CC trait of the plant. The transgenic plant is useful for altering a plant's
 CC trait for increasing plant stability. The sequence presented is a protein
 CC of the invention.
 XX
 XX Sequence 304 AA;
 SQ
 Query Match 96.6%; Score 28; DB 6; Length 304;
 Best Local Similarity 83.3%; Pred. No. 1.7e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GHXXKL 6
 DB 206 GHKAKL 211

RESULT 7
 ABU08934
 ID ABU08934 standard; protein; 407 AA.
 XX AC ABU08934;
 XX 05-JUN-2003 (first entry)
 DT Human tumour rejection antigen precursor, MAGE-B6.
 DE
 DE Human tumour rejection antigen precursor, MAGE-B6.
 XX TRAP; tumour rejection antigen precursor; cytolytic T-cell; CTL; tumour;
 KW seminoma; bladder transitional-cell carcinoma; NSCLC; adaptor;
 KW head-and-neck squamous-cell carcinoma; breast carcinoma; sarcoma;
 KW cutaneous melanoma; nonsmall cell lung cancer; MAGE-B6; human.
 XX
 OS Homo sapiens.
 XX
 XX US2002176865-A1.
 XX 28-NOV-2002.
 PD

```

XX PF 01-MAR-2002; 2002US-00085108.
XX PR 25-APR-1997; 97US-00845528.
XX PR 24-APR-1998; 98US-00062821.
XX PR 17-DEC-1999; 99US-00468433.
XX PR 09-FEB-2000; 2000US-00501104.
XX XX (LUCA/) LUCAS S.
XX PA (BOON/) BOON-FALLEUR T.
XX PI Lucas S, Boon-Falleur T;
XX PT WPI; 2003-328468/31.
XX DR N-PSDE; ABX95008.
XX XX
XX PT Novel isolated nucleic acid encoding tumor rejection antigen precursor
XX PT MAGE-C3, MAGE-B5, or MAGE-B6, useful as diagnostic probes to determine
XX PT presence of abnormal e.g., tumor cells expressing MAGE-C1, MAGE-B5 or
XX PT MAGE-B6.
XX PS Example 13; Fig 10; 59pp; English.
XX CC The invention relates to an isolated nucleic acid molecule which encodes
XX CC a tumour rejection antigen precursor (TRAP) having an amino acid sequence
XX CC of a TRAP encoded by a fully defined MAGE-C3, MAGE-B5, or MAGE-B6
XX CC polynucleotide sequence. Also disclosed is a method which is useful for
XX CC determining presence of cytolytic T-cells specific for complexes of human
XX CC leukocyte antigen (HLA) and a peptide derived from the nucleic acid in a
XX CC cytotoxic T-lymphocyte (CTL)-containing sample. The nucleic acid is
XX CC useful as a diagnostic probe to determine the presence of abnormal
XX CC (tumour) cells such as seminoma, bladder transitional-cell carcinoma,
XX CC head-and-neck squamous-cell carcinoma, breast carcinoma, sarcoma,
XX CC cutaneous melanoma or non-small cell lung cancer (NSCLC) which express
XX CC MAGE-C1, MAGE-B5 or MAGE-B6. The nucleic acid is useful for diagnosing a
XX CC disorder characterised by expression of MAGE-C1, MAGE-B5 or MAGE-B6 TRAPs
XX CC or tumour rejection antigens (TRAs). The present sequence represents the
XX CC amino acid sequence of the human tumour rejection antigen precursor, MAGE
XX CC -B6
XX SQ Sequence 407 AA;
XX
XX Query Match 36.6%; Score 28; DB 6; Length 407;
XX Best Local Similarity 83.3%; Pred. No. 2.3e+02;
XX Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 1 GHKXKL 6
XX |||||
XX 4 GHKSXL 9
XX
XX DE Peptide motif B10 targeted to human Tfr.
XX
XX RESULT 8
XX AAO30886
XX ID AAO30886 standard; peptide; 13 AA.
XX XX
XX AC AAO30886;
XX XX
XX DT 22-SEP-2003 (first entry)
XX XX
XX DE Peptide motif B10 targeted to human Tfr.
XX XX
XX KW Adenoviral fiber protein; coxsackie-adenovirus receptor; CAR; diabetes;
XX KW acquired immune deficiency syndrome; genetic disease; viral hepatitis;
XX KW gene therapy; sickle cell anaemia; Alzheimer's disease; AIDS; cancer;
XX KW haemophilia; epitope; transferrin receptor; Tfr.
XX XX
XX OS Unidentified.
XX XX
XX PN WO2003050238-A2.
XX XX
XX PD 19-JUN-2003.
XX XX
XX PF 07-NOV-2002; 2002WO-US035822.

```

```

XX 11-DEC-2001; 2001US-0339282P.
XX XX (IOWA ) UNIV IOWA RES FOUND.
XX PA Davidson BL, Xia H, Law LK;
XX PI WPI; 2003-523358/49.
XX XX
XX DR New receptor-targeted adenoviral vector comprising an adenoviral backbone
XX PT encoding an adenoviral fiber that does not bind coxsackie-adenovirus
XX PT receptor, useful for treating a genetic disease, cancer, hemophilia or
XX PT AIDS in a mammal.
XX PS Example 2; Page 65; 105pp; English.
XX XX
XX CC The invention relates to an adenoviral vector comprising an adenoviral
XX CC backbone encoding an adenoviral fiber that does not bind coxsackie-
XX CC adenovirus receptor (CAR) and an adenoviral fiber protein HI-loop
XX CC operably linked to a receptor-targeting ligand to form a ligand/HI-loop
XX CC chimeric protein, where the chimeric protein binds to a corresponding
XX CC targeted receptor but does not bind CAR. The adenoviral vector, chimeric
XX CC protein, adenovirus particle or mammalian cell is useful for treating a
XX CC genetic disease or cancer in a mammal. The invention is also useful in
XX CC gene therapy or for treating haemophilia, acquired immune deficiency
XX CC syndrome (AIDS), sickle cell anaemia, diabetes, Alzheimer's disease or
XX CC viral hepatitis. The present sequence is a peptide epitope targeted to
XX CC human transferrin receptor (Tfr). This peptide is used to illustrate the
XX CC method of the invention
XX XX
XX SQ Sequence 13 AA;
XX
XX Query Match 93.1%; Score 27; DB 6; Length 13;
XX Best Local Similarity 83.3%; Pred. No. 11;
XX Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 1 GHKXKL 6
XX |||||
XX 6 GHKXKL 11
XX
XX DE Human two-chain high molecular weight kininogen domain 5 fragment #7.
XX XX
XX KW Human; high molecular weight kininogen; HK;
XX KW two-chain high molecular weight kininogen; HKa; angiogenesis inhibition;
XX KW tumour; cancer; ocular disorder; rheumatoid arthritis;
XX KW endothelial cell apoptosis.
XX XX
XX OS Homo sapiens.
XX XX
XX PN WO200027866-A1.
XX XX
XX PD 18-MAY-2000.
XX XX
XX PF 05-NOV-1999; 99WO-US026419.
XX XX
XX PR 10-NOV-1998; 98US-0107833P.
XX XX
XX PA (UTEM ) UNIV TEMPLE.
XX PA (MCCR/) MCCR R K.
XX XX
XX PI McCrae RK;
XX XX
XX DR WPI; 2000-376483/32.
XX XX

```

PT A pharmaceutical composition used to inhibit angiogenesis, inhibit
 PT endothelial cell proliferation, and induce endothelial cell apoptosis.
 XX
 PS Claim 9; Page 28; 52pp; English.
 XX
 CC The present sequence is derived from human two-chain high molecular
 CC weight kininogen (HKa) domain 5. HKa is product of high molecular weight
 CC kininogen (HK) cleavage by plasma kallikrein. HK is a 120 kD glycoprotein
 CC which binds with high affinity to endothelial cells. HKa or a synthetic
 CC compound comprising the present sequence may be used in a pharmaceutical
 CC composition for inhibiting angiogenesis. Angiogenesis occurs in a number
 CC of disease states, such as tumour formation and expansion, and certain
 CC ocular disorders. It can also occur in a rheumatoid joint, hastening
 CC joint destruction by allowing an influx of leukocytes. The composition
 CC may inhibit angiogenesis by inhibiting endothelial cell proliferation or
 CC by inducing endothelial cell apoptosis. Peptides used in the composition
 CC may be recombinant peptides, natural peptides, or synthetic peptides.
 CC They may also be chemically synthesised, using, for example, solid phase
 CC synthesis methods
 XX
 SQ Sequence 16 AA;

Query Match 93.1%; Score 27; DB 3; Length 16;
 Best Local Similarity 83.3%; Pred. No. 14;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GHKXKL 6
 |||||
 Db 1 GHKFKL 6

RESULT 10
 ABB78700
 ID ABB78700 standard; peptide; 16 AA.
 XX
 AC ABB78700;
 XX
 DT 18-JUL-2002 (first entry)
 XX
 DE Human kininogen D5 domain peptide SEQ ID NO:4.
 XX
 KW Human; kininogen; high molecular weight kininogen; HK; D5 domain;
 KW D5 receptor; angiogenesis; endothelial cell; cytostatic; antitumour;
 KW antithrombotic; vasoregulatory; antidiabetic; antihypertensive;
 KW ophthalmological; gynaecological; antitumor; antidiabetic; antithrombotic;
 KW antiangiogenic; antiproliferative; endocrine; apoptosis; gene therapy.
 XX
 OS Homo sapiens.
 XX
 FN WO200214369-A2.
 XX
 PD 21-FEB-2002.
 XX
 PF 24-JUL-2001; 2001WO-US023185.
 XX
 PR 24-JUL-2000; 2000US-0220194P.
 XX
 PA (ATTE-) ATTENUON LLC.
 XX
 PI Mazar AP, Juarez JC;
 XX
 DR WPI; 2002-393611/42.
 XX
 XX Novel human kininogen D5 domain polypeptides useful for treating
 PT conditions associated with endothelial cell migration, proliferation,
 PT invasion or angiogenesis, e.g. arthritis, macular degeneration, benign
 PT hyperplasia.
 XX
 PS Claim 3; Page 65; 84pp; English.
 XX
 CC The present invention describes an isolated polypeptide (I) that
 CC corresponds to the D5 domain of human kininogen or biologically active
 CC peptide fragment, homologue or functional derivative, and which: (a)

CC inhibits angiogenesis; (b) binds to the D5 binding site on endothelial
 CC cells (EC); (c) activates signalling pathways leading to the introduction
 CC of apoptosis in EC; and/or (d) inhibits the signalling pathway required
 CC for maintenance of EC viability. (I) has cytostatic, antitumour,
 CC antithrombotic, vasoregulatory, antidiabetic, antihypertensive, thrombolytic,
 CC ophthalmological, gynaecological, antitumor, antidiabetic, antithrombotic,
 CC antiangiogenic, antiproliferative and endocrine activities. An antibody (IX)
 CC specific for an epitope of (I) is useful for inhibiting tumour growth or
 CC angiogenesis in a subject. (I), a D5 fusion polypeptide (II) or a dimeric
 CC or trimeric fusion polypeptide (III) can be used for inhibiting EC
 CC migration, proliferation, invasion, or angiogenesis, or for inducing EC
 CC apoptosis. An angiogenic EC-targeting pharmaceutical composition (X)
 CC comprising (I), (II), or (III), can be used for treating a subject having
 CC a disease or condition associated with undesired EC migration,
 CC proliferation, invasion or angiogenesis. (I), (II), or (III) can be used
 CC for isolating a D5 domain binding molecule from a complex mixture and for
 CC isolating or enriching cells expressing D5 domain binding sites from a
 CC cell mixture. The present sequence represents a human high molecular
 CC weight kininogen (HK) D5 domain peptide from the present invention
 XX
 SQ Sequence 16 AA;

Query Match 93.1%; Score 27; DB 5; Length 16;
 Best Local Similarity 83.3%; Pred. No. 14;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GHKXKL 6
 |||||
 Db 1 GHKFKL 6

RESULT 11
 AAY81996
 ID AAY81996 standard; peptide; 28 AA.
 XX
 AC AAY81996;
 XX
 DT 16-OCT-2000 (first entry)
 XX
 DE Human high molecular weight kininogen domain 5 fragment #5.
 XX
 KW Human; high molecular weight kininogen; HK;
 KW two-chain high molecular weight kininogen; HKa; angiogenesis inhibition;
 KW tumour; cancer; ocular disorder; rheumatoid arthritis;
 KW endothelial cell apoptosis.
 XX
 OS Homo sapiens.
 XX
 FN WO200027866-A1.
 XX
 PD 18-MAY-2000.
 XX
 PF 05-NOV-1999; 99WO-US026419.
 XX
 PR 10-NOV-1998; 98US-0107833P.
 XX
 PA (UTEM) UNIV TEMPLE.
 XX
 PI (MCCR/) MCCRAE R K.
 XX
 DR McCrae RK;
 XX
 DR WPI; 2000-376483/32.
 XX
 PT A pharmaceutical composition used to inhibit angiogenesis, inhibit
 PT endothelial cell proliferation, and induce endothelial cell apoptosis.
 XX
 PS Claim 8; Page 28; 52pp; English.
 XX
 CC The present sequence is derived from human high molecular weight
 CC kininogen (HK) domain 5. HK is a 120 kD glycoprotein which binds with
 CC high affinity to endothelial cells, where it is cleaved to two-chain high
 CC molecular weight kininogen (HKa) by plasma kallikrein. HKa or a synthetic
 CC compound comprising the present sequence may be used in a pharmaceutical

CC composition for inhibiting angiogenesis. Angiogenesis occurs in a number
 CC of disease states, such as tumour formation and expansion, and certain
 CC ocular disorders. It can also occur in a rheumatoid joint, hastening
 CC joint destruction by allowing an influx of leukocytes. The composition
 CC may inhibit angiogenesis by inhibiting endothelial cell proliferation or
 CC by inducing endothelial cell apoptosis. Peptides used in the composition
 CC may be recombinant peptides, natural peptides, or synthetic peptides.
 CC They may also be chemically synthesised, using, for example, solid phase
 CC synthesis methods
 XX
 SQ Sequence 28 AA;

Query Match 93.1%; Score 27; DB 3; Length 28;
 Best Local Similarity 83.3%; Pred. No. 24;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GHKXKL 6
 DB 12 GHKFKL 17
 |||||

RESULT 12

AA93345
 ID AAY93345 standard; peptide; 47 AA.

AC AAY93345;

DT 04-SEP-2000 (first entry)

DE Light chain of human high molecular weight kininogen fragment.

XX Human; high molecular weight kininogen; glycoprotein; endothelial cell;
 KW plasma kallikrein; heavy chain; light chain; analogue; angiogenesis;
 KW endothelial cell proliferation; endothelial cell migration; vitronectin.

XX Synthetic.

OS Homo sapiens.

XX WO200027415-A2.

XX 18-MAY-2000.

XX 09-NOV-1999; 99WO-US026377.

XX 10-NOV-1998; 98US-0107844P.

XX (UTEM) UNIV TEMPLE.

PA (DUPO) DUPONT PHARM CO.

PA (COLM/) COLMAN W R.

PA (MOUS/) MOUSA A S.

XX Colman WR, Mousa AS;

XX WPI; 2000-376306/32.

XX Claim 3; Page 36; 41pp; English.

CC The present sequence represents a fragment of the light chain of human
 CC high molecular weight kininogen. It is used to produce compounds of the
 CC invention. High molecular weight kininogen is a 120 kDa glycoprotein
 CC which binds with high affinity to endothelial cells, where it is cleaved
 CC by plasma kallikrein into heavy and light chains. Analogues of high
 CC molecular weight kininogen are used in the method of the invention. The
 CC specification describes a method of inhibiting endothelial cell
 CC proliferation. The method comprises contacting endothelial cells with a
 CC compound containing high molecular weight kininogen analogues. The method
 CC and the compounds can be used for inhibiting endothelial cell
 CC proliferation. The compounds can also be used for inhibiting
 CC angiogenesis. The compounds can also be used to inhibit migration of
 CC endothelial cells to vitronectin

XX Sequence 47 AA;
 SQ

Query Match 93.1%; Score 27; DB 3; Length 47;
 Best Local Similarity 83.3%; Pred. No. 42;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GHKXKL 6
 DB 1 GHKFKL 6
 |||||

RESULT 13

AA93346
 ID AAY93346 standard; peptide; 55 AA.

AC AAY93346;

DT 04-SEP-2000 (first entry)

DE Light chain of human high molecular weight kininogen analogue.

XX Human; high molecular weight kininogen; glycoprotein; endothelial cell;
 KW plasma kallikrein; heavy chain; light chain; analogue; angiogenesis;
 KW endothelial cell proliferation; endothelial cell migration; vitronectin.

XX Synthetic.

OS Homo sapiens.

XX WO200027415-A2.

XX 18-MAY-2000.

XX 09-NOV-1999; 99WO-US026377.

XX 10-NOV-1998; 98US-0107844P.

XX (UTEM) UNIV TEMPLE.

PA (DUPO) DUPONT PHARM CO.

PA (COLM/) COLMAN W R.

PA (MOUS/) MOUSA A S.

XX Colman WR, Mousa AS;

XX WPI; 2000-376306/32.

XX Method for inhibiting endothelial cell proliferation, using compound that
 PT inhibit endothelial cell migration.

XX Claim 4; Page 36; 41pp; English.

CC The present sequence represents an analogue of the light chain of human
 CC high molecular weight kininogen. High molecular weight kininogen is a 120
 CC kDa glycoprotein which binds with high affinity to endothelial cells,
 CC where it is cleaved by plasma kallikrein into heavy and light chains.
 CC Analogues of high molecular weight kininogen are used in the method of
 CC the invention. The specification describes a method of inhibiting
 CC endothelial cell proliferation. The method comprises contacting
 CC endothelial cells with a compound containing high molecular weight
 CC kininogen analogues. The method and the compounds can be used for
 CC inhibiting endothelial cell proliferation. The compounds can also be used
 CC for inhibiting angiogenesis. The compounds can also be used to inhibit
 CC migration of endothelial cells to vitronectin

XX Sequence 55 AA;

Query Match 93.1%; Score 27; DB 3; Length 55;
 Best Local Similarity 83.3%; Pred. No. 49;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GHKXKL 6
 DB 37 GHKFKL 42
 |||||

RESULT 14
 AAY93348
 ID AAY93348 standard; peptide; 62 AA.
 AC AAY93348;
 DT 04-SEP-2000 (first entry)
 DE Light chain of human high molecular weight kininogen analogue.
 XX Human; high molecular weight kininogen; glycoprotein; endothelial cell;
 KW plasma kallikrein; heavy chain; light chain; analogue; angiogenesis;
 KW endothelial cell proliferation; endothelial cell migration; vitronectin.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 PN WO200027415-A2.
 XX
 PD 18-MAY-2000.
 XX
 PF 09-NOV-1999; 99WO-US026377.
 XX
 PR 10-NOV-1998; 98US-0107844P.
 XX
 PA (UTEM) UNIV TEMPLE.
 PA (DUPO) DUFONT PHARM CO.
 PA (COLM) COLMAN W R.
 PA (MOUS) MOUSA A S.
 XX
 PI Colman WR, Mousa AS;
 XX
 DR WPI; 2000-376306/32.
 XX
 PT Method for inhibiting endothelial cell proliferation, using compound that
 PT inhibit endothelial cell migration.
 XX
 PS Claim 6; Page 37; 41pp; English.
 XX
 CC The present sequence represents an analogue of the light chain of human
 CC high molecular weight kininogen. High molecular weight kininogen is a 120
 CC kDa glycoprotein which binds with high affinity to endothelial cells,
 CC where it is cleaved by plasma kallikrein into heavy and light chains.
 CC Analogues of high molecular weight kininogen are used in the method of
 CC the invention. The specification describes a method of inhibiting
 CC endothelial cell proliferation. The method comprises contacting
 CC endothelial cells with a compound containing high molecular weight
 CC kininogen analogues. The method and the compounds can be used for
 CC inhibiting endothelial cell proliferation. The compounds can also be used
 CC for inhibiting angiogenesis. The compounds can also be used to inhibit
 CC migration of endothelial cells to vitronectin
 XX
 SQ Sequence 62 AA;
 Query Match 93.1%; Score 27; DB 3; Length 62;
 Best Local Similarity 83.3%; Pred. No. 56;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 GHXXKL 6
 Db 16 GHKFKL 21
 RESULT 15
 AAR75186
 ID AAR75186 standard; peptide; 63 AA.
 AC AAR75186;
 XX
 DT 05-DEC-1995 (first entry)
 XX

DE Partial peptide of human HMW kininogen fragment 2.
 XX high molecular weight; kininogen; fragment; 1.2; 1; 2; partial;
 KW wound treating agent; bovine; growth promotion; fibroblast.
 XX
 OS Homo sapiens.
 PN JP07082172-A.
 XX
 PD 28-MAR-1995.
 XX
 PF 17-SEP-1993; 93JP-00230616.
 XX
 PR 17-SEP-1993; 93JP-00230616.
 XX
 PA (PARH) HOECHST JAPAN KK.
 XX
 DR WPI; 1995-158909/21.
 XX
 PT A wound treating agent contg. a partial peptide of kininogen - have
 PT growth promotion activity of fibroblasts.
 XX
 PS Claim 8; Page 8; 8pp; Japanese.
 XX
 CC AAR75186 is a partial peptide corresponding to human kininogen fragment
 CC 1, amino acids 459-520. Partial peptides of bovine and human kininogen
 CC fragments 1.2, 1 and 2, are used in wound treating agent compans. and act
 CC as the active component. The fragments are useful in wound treating
 CC because they have growth promotion activity on fibroblasts
 XX
 SQ Sequence 63 AA;
 Query Match 93.1%; Score 27; DB 2; Length 63;
 Best Local Similarity 83.3%; Pred. No. 57;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 GHXXKL 6
 Db 17 GHKFKL 22
 Search completed: March 4, 2004, 13:07:29
 Job time : 18.6765 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2004, 13:04:08 ; Search time 4.58824 Seconds
(without alignments)
67.511 Million cell updates/sec

Title: SEQ-D
Perfect score: 29
Sequence: 1 ghkxkl 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents.MA*
1: /cgn2_6/ptodata/2/iaa/5A COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PGTUS COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	28	96.6	59	3	US-09-300-672-7
2	28	96.6	407	4	US-09-468-433C-26
3	27	93.1	47	3	US-09-612-126-4
4	27	93.1	55	3	US-09-612-126-5
5	27	93.1	62	3	US-09-612-126-7
6	27	93.1	83	3	US-09-612-126-6
7	27	93.1	94	3	US-09-612-126-10
8	27	93.1	179	3	US-09-612-126-11
9	27	93.1	186	3	US-09-612-126-8
10	27	93.1	249	4	US-09-252-991A-32233
11	27	93.1	255	3	US-09-612-126-1
12	27	93.1	509	1	US-08-030-096-2
13	27	93.1	632	4	US-09-252-991A-17148
14	27	93.1	844	4	US-09-252-991A-27184
15	25	86.2	265	4	US-09-543-681A-7082
16	25	86.2	300	4	US-09-976-594-7
17	25	86.2	371	4	US-09-543-681A-4389
18	25	86.2	409	4	US-09-543-681A-4445
19	25	86.2	456	4	US-09-252-991A-28041
20	25	86.2	993	4	US-09-894-998A-50
21	25	86.2	1037	4	US-09-894-998A-54
22	25	86.2	1113	4	US-09-894-998A-51
23	24	82.8	7	4	US-09-627-851B-48
24	24	82.8	32	1	US-08-190-802A-111
25	24	82.8	32	1	US-08-190-802A-115
26	24	82.8	32	3	US-08-477-346-111
27	24	82.8	32	3	US-08-477-346-115

Sequence 111, App
Sequence 115, App
Sequence 111, App
Sequence 115, App
Sequence 5354, App
Sequence 7, Appli
Sequence 14303, A
Sequence 7927, Ap
Sequence 4, Appli
Sequence 5248, Ap
Sequence 6, Appli
Sequence 31936, A
Sequence 4390, Ap
Sequence 51, Appl
Sequence 51, Appl
Sequence 51, Appl
Sequence 51, Appl
Sequence 10472, A

ALIGNMENTS

RESULT 1
US-09-300-672-7
; Sequence 7, Application US/09300672
; Patent No. 6248937
; GENERAL INFORMATION:
; APPLICANT: Finkelstein, Ruth R.
; APPLICANT: Lynch, Tim
; APPLICANT: Goodman, Howard M.
; APPLICANT: Wang, Ming-Li
; TITLE OF INVENTION: A TRANSCRIPTION FACTOR REGULATING SEED DEVELOPMENT,
; TITLE OF INVENTION: QUALITY AND COLD-TOLERANCE
; FILE REFERENCE: 480.85(HV)
; CURRENT APPLICATION NUMBER: US/09/300,672
; CURRENT FILING DATE: 1999-04-27
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 59
; TYPE: PRT
; ORGANISM: AP2 domain protein
US-09-300-672-7

Query Match 96.6%; Score 28; DB 3; Length 59;
Best Local Similarity 83.3%; Pred. No. 11;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GHXXKL 6
|||
DB 49 GHKAKL 54

RESULT 2
US-09-468-433C-26
; Sequence 26, Application US/09468433C
; Patent No. 6680191
; GENERAL INFORMATION:
; APPLICANT: LUCAS, Sophie; BOON-FALLEUR, Thierry
; TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULES CODING FOR
; TITLE OF INVENTION: TUMOR REJECTION ANTIGEN PRECURSORS OF MEMBERS OF THE MAGE-C ANI
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fulbright & Jaworski L.L.P.
; STREET: 801 Pennsylvania Avenue, NW
; CITY: Washington
; STATE: District of Columbia
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage

; COMPUTER: IBM PS/2
 ; OPERATING SYSTEM: PC-DOS
 ; SOFTWARE: Wordperfect
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/468,433C
 ; FILING DATE: December 17, 1999
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 09/066,281
 ; FILING DATE: April 24, 1998
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/845,528
 ; FILING DATE: April 25, 1997
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Mary Anne Schofield
 ; REGISTRATION NUMBER: 36,669
 ; REFERENCE/DOCKET NUMBER: LUD 5611 JEL/MAS
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (202) 662-0200
 ; TELEFAX: (202) 662-4643
 ; INFORMATION FOR SEQ ID NO: 26:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 407
 ; TYPE: amino acid
 ; STRANDEDNESS: single-stranded
 ; TOPOLOGY: linear
 ; US-09-468-433C-26

Query Match 96.6%; Score 28; DB 3; Length 407;
 Best Local Similarity 83.3%; Pred. No. 72;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GHXXKL 6
 |||||
 Db 4 GHKSXL 9

RESULT 3
 US-09-612-126-4
 ; Sequence 4, Application US/09612126
 ; Patent No. 6284726
 ; GENERAL INFORMATION:
 ; APPLICANT: TEMPLE UNIVERSITY - OF THE COMMONWEALTH SYSTEM OF
 ; TITLE OF INVENTION: INHIBITION OF ANGIOGENESIS BY PEPTIDE ANALOGS OF HIGH
 ; FILE REFERENCE: 6056-258 CT1
 ; CURRENT APPLICATION NUMBER: US/09/612,126
 ; CURRENT FILING DATE: 2000-07-07
 ; PRIOR APPLICATION NUMBER: 60/107,844
 ; PRIOR FILING DATE: 1998-11-10
 ; PRIOR APPLICATION NUMBER: PCT/US99/26377
 ; PRIOR FILING DATE: 1999-11-09
 ; NUMBER OF SEQ ID NOS: 12
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 4
 ; LENGTH: 47
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: Human high
 ; OTHER INFORMATION: molecular weight kininogen light chain amino acids
 ; OTHER INFORMATION: Gly(456) through Lys(502)
 ; US-09-612-126-4

Query Match 93.1%; Score 27; DB 3; Length 47;
 Best Local Similarity 83.3%; Pred. No. 15;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GHXXKL 6
 |||||
 Db 1 GHKFKL 6

RESULT 4
 US-09-612-126-5
 ; Sequence 5, Application US/09612126
 ; Patent No. 6284726
 ; GENERAL INFORMATION:
 ; APPLICANT: TEMPLE UNIVERSITY - OF THE COMMONWEALTH SYSTEM OF
 ; TITLE OF INVENTION: INHIBITION OF ANGIOGENESIS BY PEPTIDE ANALOGS OF HIGH
 ; FILE REFERENCE: 6056-258 CT1
 ; CURRENT APPLICATION NUMBER: US/09/612,126
 ; CURRENT FILING DATE: 2000-07-07
 ; PRIOR APPLICATION NUMBER: 60/107,844
 ; PRIOR FILING DATE: 1998-11-10
 ; PRIOR APPLICATION NUMBER: PCT/US99/26377
 ; PRIOR FILING DATE: 1999-11-09
 ; NUMBER OF SEQ ID NOS: 12
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 5
 ; LENGTH: 55
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: Human high
 ; OTHER INFORMATION: molecular weight kininogen light chain amino acids
 ; OTHER INFORMATION: Lys(420) through Asp(474)
 ; US-09-612-126-5

Query Match 93.1%; Score 27; DB 3; Length 55;
 Best Local Similarity 83.3%; Pred. No. 17;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GHXXKL 6
 |||||
 Db 37 GHKFKL 42

RESULT 5
 US-09-612-126-7
 ; Sequence 7, Application US/09612126
 ; Patent No. 6284726
 ; GENERAL INFORMATION:
 ; APPLICANT: TEMPLE UNIVERSITY - OF THE COMMONWEALTH SYSTEM OF
 ; TITLE OF INVENTION: INHIBITION OF ANGIOGENESIS BY PEPTIDE ANALOGS OF HIGH
 ; FILE REFERENCE: 6056-258 CT1
 ; CURRENT APPLICATION NUMBER: US/09/612,126
 ; CURRENT FILING DATE: 2000-07-07
 ; PRIOR APPLICATION NUMBER: 60/107,844
 ; PRIOR FILING DATE: 1998-11-10
 ; PRIOR APPLICATION NUMBER: PCT/US99/26377
 ; PRIOR FILING DATE: 1999-11-09
 ; NUMBER OF SEQ ID NOS: 12
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 7
 ; LENGTH: 62
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: Human high
 ; OTHER INFORMATION: molecular weight kininogen light chain amino acids
 ; OTHER INFORMATION: His(441) through Lys(502)
 ; US-09-612-126-7

Query Match 93.1%; Score 27; DB 3; Length 62;
 Best Local Similarity 83.3%; Pred. No. 19;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GHXXKL 6
 |||||
 Db 16 GHKFKL 21

RESULT 6

US-09-612-126-6
 ; Sequence 6, Application US/09612126
 ; Patent No. 6284726
 ; GENERAL INFORMATION:
 ; APPLICANT: TEMPLE UNIVERSITY - OF THE COMMONWEALTH SYSTEM OF
 ; TITLE OF INVENTION: INHIBITION OF ANGIOGENESIS BY PEPTIDE ANALOGS OF HIGH
 ; TITLE OF INVENTION: MOLECULAR WEIGHT KININOGEN DOMAN 5
 ; FILE REFERENCE: 6056-258 CTI
 ; CURRENT APPLICATION NUMBER: US/09/612,126
 ; CURRENT FILING DATE: 2000-07-07
 ; PRIOR APPLICATION NUMBER: 60/107,844
 ; PRIOR FILING DATE: 1998-11-10
 ; PRIOR APPLICATION NUMBER: PCT/US99/26377
 ; PRIOR FILING DATE: 1999-11-09
 ; NUMBER OF SEQ ID NOS: 12
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 6
 ; LENGTH: 83
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: Human high
 ; OTHER INFORMATION: molecular weight kininogen light chain amino acids
 ; OTHER INFORMATION: Lys(420) through Lys(502)
 US-09-612-126-6

Query Match 93.1%; Score 27; DB 3; Length 83;
 Best Local Similarity 83.3%; Pred. No. 26;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GHXXKL 6
 |||||
 Db 37 GHXFKL 42

RESULT 7
 US-09-612-126-10
 ; Sequence 10, Application US/09612126
 ; Patent No. 6284726
 ; GENERAL INFORMATION:
 ; APPLICANT: TEMPLE UNIVERSITY - OF THE COMMONWEALTH SYSTEM OF
 ; TITLE OF INVENTION: INHIBITION OF ANGIOGENESIS BY PEPTIDE ANALOGS OF HIGH
 ; TITLE OF INVENTION: MOLECULAR WEIGHT KININOGEN DOMAN 5
 ; FILE REFERENCE: 6056-258 CTI
 ; CURRENT APPLICATION NUMBER: US/09/612,126
 ; CURRENT FILING DATE: 2000-07-07
 ; PRIOR APPLICATION NUMBER: 60/107,844
 ; PRIOR FILING DATE: 1998-11-10
 ; PRIOR APPLICATION NUMBER: PCT/US99/26377
 ; PRIOR FILING DATE: 1999-11-09
 ; NUMBER OF SEQ ID NOS: 12
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 10
 ; LENGTH: 94
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: Human high
 ; OTHER INFORMATION: molecular weight kininogen light chain amino acids
 ; OTHER INFORMATION: Lys(420) through Ser(513)
 US-09-612-126-10

Query Match 93.1%; Score 27; DB 3; Length 94;
 Best Local Similarity 83.3%; Pred. No. 29;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GHXXKL 6
 |||||
 Db 37 GHXFKL 42

RESULT 8
 US-09-612-126-11

; Sequence 11, Application US/09612126
 ; Patent No. 6284726
 ; GENERAL INFORMATION:
 ; APPLICANT: TEMPLE UNIVERSITY - OF THE COMMONWEALTH SYSTEM OF
 ; TITLE OF INVENTION: INHIBITION OF ANGIOGENESIS BY PEPTIDE ANALOGS OF HIGH
 ; TITLE OF INVENTION: MOLECULAR WEIGHT KININOGEN DOMAN 5
 ; FILE REFERENCE: 6056-258 CTI
 ; CURRENT APPLICATION NUMBER: US/09/612,126
 ; CURRENT FILING DATE: 2000-07-07
 ; PRIOR APPLICATION NUMBER: 60/107,844
 ; PRIOR FILING DATE: 1998-11-10
 ; PRIOR APPLICATION NUMBER: PCT/US99/26377
 ; PRIOR FILING DATE: 1999-11-09
 ; NUMBER OF SEQ ID NOS: 12
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 11
 ; LENGTH: 179
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: Human high
 ; OTHER INFORMATION: molecular weight kininogen light chain amino acids
 ; OTHER INFORMATION: Glu(448) through Ser(626)
 US-09-612-126-11

Query Match 93.1%; Score 27; DB 3; Length 179;
 Best Local Similarity 83.3%; Pred. No. 55;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GHXXKL 6
 |||||
 Db 9 GHXFKL 14

RESULT 9
 US-09-612-126-8
 ; Sequence 8, Application US/09612126
 ; Patent No. 6284726
 ; GENERAL INFORMATION:
 ; APPLICANT: TEMPLE UNIVERSITY - OF THE COMMONWEALTH SYSTEM OF
 ; TITLE OF INVENTION: INHIBITION OF ANGIOGENESIS BY PEPTIDE ANALOGS OF HIGH
 ; TITLE OF INVENTION: MOLECULAR WEIGHT KININOGEN DOMAN 5
 ; FILE REFERENCE: 6056-258 CTI
 ; CURRENT APPLICATION NUMBER: US/09/612,126
 ; CURRENT FILING DATE: 2000-07-07
 ; PRIOR APPLICATION NUMBER: 60/107,844
 ; PRIOR FILING DATE: 1998-11-10
 ; PRIOR APPLICATION NUMBER: PCT/US99/26377
 ; PRIOR FILING DATE: 1999-11-09
 ; NUMBER OF SEQ ID NOS: 12
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 8
 ; LENGTH: 186
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: Human high
 ; OTHER INFORMATION: molecular weight kininogen light chain amino acids
 ; OTHER INFORMATION: His(441) through Ser(626)
 US-09-612-126-8

Query Match 93.1%; Score 27; DB 3; Length 186;
 Best Local Similarity 83.3%; Pred. No. 57;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GHXXKL 6
 |||||
 Db 16 GHXFKL 21

RESULT 10
 US-09-252-991A-32233
 ; Sequence 32233, Application US/09252991A

```
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 32233
; LENGTH: 249
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-32233

Query Match          93.1%; Score 27; DB 4; Length 249;
Best Local Similarity 83.3%; Pred. No. 76;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GHXXKL 6
Db 188 GHXXKL 193

RESULT 11
US-09-612-126-1
; Sequence 1, Application US/09612126
; Patent No. 6284726
; GENERAL INFORMATION:
; APPLICANT: TEMPLE UNIVERSITY - OF THE COMMONWEALTH SYSTEM OF
; TITLE OF INVENTION: INHIBITION OF ANGIOGENESIS BY PEPTIDE ANALOGS OF HIGH
; FILE REFERENCE: 6056-258 Ctl
; CURRENT APPLICATION NUMBER: US/09/612,126
; CURRENT FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/107,844
; PRIOR FILING DATE: 1998-11-10
; PRIOR APPLICATION NUMBER: PCT/US99/26377
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 255
; TYPE: PRT
; ORGANISM: Human
US-09-612-126-1

Query Match          93.1%; Score 27; DB 3; Length 255;
Best Local Similarity 83.3%; Pred. No. 78;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GHXXKL 6
Db 85 GHXXKL 90

RESULT 12
US-08-030-096-2
; Sequence 2, Application US/08030096
; Patent No. 5426041
; GENERAL INFORMATION:
; APPLICANT: Rabinjanski, Steven P.
; APPLICANT: Ardison, Paul G.
; TITLE OF INVENTION: BINARY CRYPTOCYTOTOXIC METHOD OF HYBRID
; TITLE OF INVENTION: SEED PRODUCTION
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
```

```
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/030,096
; FILING DATE: 22-MAR-1993
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/556,917
; FILING DATE: 20-JUL-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/CA91/00255
; FILING DATE: 22-JUL-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 33229/164/PIHI
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 509 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-030-096-2

Query Match          93.1%; Score 27; DB 1; Length 509;
Best Local Similarity 83.3%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GHXXKL 6
Db 193 GHXXKL 198

RESULT 13
US-09-252-991A-17148
; Sequence 17148, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 17148
; LENGTH: 632
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-17148

Query Match          93.1%; Score 27; DB 4; Length 632;
Best Local Similarity 83.3%; Pred. No. 1.9e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GHXXKL 6
Db 419 GHXXKL 424
```

RESULT 14
US-09-252-991A-27184
; Sequence 27184, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 27184
; LENGTH: 844
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-27184

Query Match 93.1%; Score 27; DB 4; Length 844;
Best Local Similarity 83.3%; Pred. No. 2.5e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GHXXKL 6
||| ||
Db 317 GHVKL 322

RESULT 15
US-09-543-681A-7082
; Sequence 7082, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 7082
; LENGTH: 265
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-7082

Query Match 86.2%; Score 25; DB 4; Length 265;
Best Local Similarity 66.7%; Pred. No. 2.3e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 GHXXKL 6
||| |:
Db 138 GHGKI 143

Search completed: March 4, 2004, 13:12:27
JOB time : 5.58824 secs

Query Match 96.6%; Score 28; DB 14;
Best Local Similarity 83.38;
Best Overall Similarity 83.38;
Best NCBI Similarity 83.38;
Best NCBI Score 28; DB 14;

Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GHXXKL 6
Db 118 GHKAKL 123

RESULT 2

US-10-374-780A-2216
; Sequence 2216, Application US/10374780A
; Publication No. US20040019927A1

GENERAL INFORMATION:

APPLICANT: Sherman, Bradley K
APPLICANT: Riechmann, Jose Luis
APPLICANT: Jiang, Cai-Zhong
APPLICANT: Heard, Jacqueline E
APPLICANT: Haake, Volker
APPLICANT: Cressman, Robert A
APPLICANT: Ratcliffe, Oliver
APPLICANT: Adam, Luc J
APPLICANT: Reuber, T. Lynne
APPLICANT: Keddie, James
APPLICANT: Broun, Pierre E
APPLICANT: Pilgrim, Marsha L
APPLICANT: Dubell III, Arnold T
APPLICANT: Pineda, Omaira
APPLICANT: Yu, Guo-Liang

TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES IN PLANTS

FILE REFERENCE: MBI-0047 CIP
CURRENT APPLICATION NUMBER: US/10/374,780A

CURRENT FILING DATE: 2003-02-25

PRIOR APPLICATION NUMBER: 09/837,944

PRIOR FILING DATE: 2001-04-18

PRIOR APPLICATION NUMBER: 60/310,847

PRIOR FILING DATE: 2001-08-09

PRIOR APPLICATION NUMBER: 09/934,455

PRIOR FILING DATE: 2001-08-22

PRIOR APPLICATION NUMBER: 60/336,049

PRIOR FILING DATE: 2001-11-19

PRIOR APPLICATION NUMBER: 60/338,692

PRIOR FILING DATE: 2001-12-11

PRIOR APPLICATION NUMBER: 10/171,468

PRIOR FILING DATE: 2002-06-14

PRIOR APPLICATION NUMBER: 10/225,066

PRIOR FILING DATE: 2002-08-09

PRIOR APPLICATION NUMBER: 10/225,067

PRIOR FILING DATE: 2002-08-09

PRIOR APPLICATION NUMBER: 10/225,068

PRIOR FILING DATE: 2002-08-09

NUMBER OF SEQ ID NOS: 2906

SOFTWARE: PatentIn version 3.2

SEQ ID NO 2216

LENGTH: 218

TYPE: PRT

FEATURE:

ORGANISM: Arabidopsis thaliana

OTHER INFORMATION: G26

US-10-374-780A-2216

Query Match 96.6%; Score 28; DB 15; Length 218;
Best Local Similarity 83.3%; Pred. No. 1.5e+02;

Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GHXXKL 6
Db 118 GHKAKL 123

RESULT 3

US-10-278-173-30

; Sequence 30, Application US/10278173

; Publication No. US20030061637A1

GENERAL INFORMATION:

APPLICANT: Jiang, Cai-Zhong
APPLICANT: Broun, Pierre
APPLICANT: Riechmann, Jose-Luis
APPLICANT: Pineda, Omaira
APPLICANT: Zhang, James
APPLICANT: Yu, Guo-Liang
APPLICANT: Pilgrim, Marsha
APPLICANT: Keddie, James
APPLICANT: Heard, Jacqueline
APPLICANT: Reuber, Lynne
APPLICANT: Ratcliffe, Oliver
APPLICANT: Adam, Luc
APPLICANT: Samaha, Raymond

TITLE OF INVENTION: POLYNUCLEOTIDES FOR ROOT TRAIT ALTERATION

FILE REFERENCE: MBI-009

CURRENT APPLICATION NUMBER: US/10/278,173

CURRENT FILING DATE: 2002-10-21

PRIOR APPLICATION NUMBER: US/09/533,392

PRIOR FILING DATE: 2000-03-22

PRIOR APPLICATION NUMBER: 60/125,814

PRIOR FILING DATE: 1999-03-23

NUMBER OF SEQ ID NOS: 177

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 30

LENGTH: 304

TYPE: PRT

ORGANISM: Arabidopsis thaliana

FEATURE:

OTHER INFORMATION: G1474

US-10-278-173-30

Query Match 96.6%; Score 28; DB 14; Length 304;
Best Local Similarity 83.3%; Pred. No. 2e+02;

Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GHXXKL 6
Db 206 GHKAKL 211

RESULT 4

US-10-085-108-26

; Sequence 26, Application US/10085108

; Publication No. US2002017685A1

GENERAL INFORMATION:

APPLICANT: LUCAS, Sophie; BOON-FALLEUR, Thierry

TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULES CODING FOR TUMOR REJECTION ANTIGEN PRECURSORS OF MEMBERS OF THE MAGE-C

MAGE-B FAMILIES AND USES THEREOF

NUMBER OF SEQUENCES: 26

CORRESPONDENCE ADDRESS:

ADDRESSEE: Fulbright & Jaworski L.L.P.

STREET: 666 Fifth Avenue

CITY: New York City

STATE: New York

COUNTRY: USA

ZIP: 10103

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage

COMPUTER: IBM PS/2

OPERATING SYSTEM: PC-DOS

SOFTWARE: Wordperfect

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/085,108

FILING DATE: 01-Mar-2002

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/501,104

FILING DATE: 09-Feb-2000

APPLICATION NUMBER: 09/468,433

FILING DATE: December 17, 1999

APPLICATION NUMBER: 09/066,281

```
/ FILING DATE: April 24, 1998
/ APPLICATION NUMBER: 08/845,528
/ FILING DATE: April 25, 1997
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Mary Anne Schofield
/ REGISTRATION NUMBER: 36,669
/ REFERENCE/DOCKET NUMBER: LUD 5611.1 JEL/WAS
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (212) 318-3100
/ TELEFAX: (212) 318-3400
/ INFORMATION FOR SEQ ID NO: 26:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 407
/ TYPE: amino acid
/ STRANDEDNESS: single-stranded
/ TOPOLOGY: linear
/ SEQUENCE DESCRIPTION: SEQ ID NO: 26:
US-10-085-108-26

Query Match          96.6%; Score 28; DB 13; Length 407;
Best Local Similarity 83.3%; Pred. No. 2.7e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GHXXKL 6
Db 4 GHXXKL 9

RESULT 5
US-10-108-260A-4444
/ Sequence 4444, Application US/10108260A
/ Publication No. US20040005560A1
/ GENERAL INFORMATION:
/ APPLICANT: HELIX RESEARCH INSTITUTE
/ TITLE OF INVENTION: No. US20040005560A1e1 full length cDNA
/ FILE REFERENCE: HQ-A0106
/ CURRENT APPLICATION NUMBER: US/10/108,260A
/ CURRENT FILING DATE: 2002-03-27
/ NUMBER OF SEQ ID NOS: 5458
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 4444
/ LENGTH: 407
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-108-260A-4444

Query Match          96.6%; Score 28; DB 15; Length 407;
Best Local Similarity 83.3%; Pred. No. 2.7e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GHXXKL 6
Db 4 GHXXKL 9

RESULT 6
US-10-291-250-33
/ Sequence 33, Application US/10291250
/ Publication No. US20030133912A1
/ GENERAL INFORMATION:
/ APPLICANT: University of Iowa Research Foundation
/ APPLICANT: Davidson, Beverly L.
/ APPLICANT: Xia, Haibin
/ APPLICANT: Law, Jane K.
/ TITLE OF INVENTION: RECEPTOR-TARGETED ADENOVIRAL VECTORS
/ FILE REFERENCE: 875,051US1
/ CURRENT APPLICATION NUMBER: US/10/291,250
/ CURRENT FILING DATE: 2002-11-07
/ PRIOR APPLICATION NUMBER: US 60/339,282
/ PRIOR FILING DATE: 2001-12-11
/ NUMBER OF SEQ ID NOS: 40
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 33

/ FILING DATE: April 24, 1998
/ APPLICATION NUMBER: 08/845,528
/ FILING DATE: April 25, 1997
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Mary Anne Schofield
/ REGISTRATION NUMBER: 36,669
/ REFERENCE/DOCKET NUMBER: LUD 5611.1 JEL/WAS
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (212) 318-3100
/ TELEFAX: (212) 318-3400
/ INFORMATION FOR SEQ ID NO: 27:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 407
/ TYPE: amino acid
/ STRANDEDNESS: single-stranded
/ TOPOLOGY: linear
/ SEQUENCE DESCRIPTION: SEQ ID NO: 27:
US-10-085-108-27

Query Match          96.6%; Score 28; DB 13; Length 407;
Best Local Similarity 83.3%; Pred. No. 2.7e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GHXXKL 6
Db 4 GHXXKL 9

RESULT 7
US-10-106-698-5176
/ Sequence 5176, Application US/10106698
/ Publication No. US20030109690A1
/ GENERAL INFORMATION:
/ APPLICANT: Ruben et al.
/ TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptides
/ FILE REFERENCE: PA005P1
/ CURRENT APPLICATION NUMBER: US/10/106,698
/ CURRENT FILING DATE: 2002-03-27
/ PRIOR APPLICATION NUMBER: PCT/US00/26524
/ PRIOR FILING DATE: 2000-09-28
/ PRIOR APPLICATION NUMBER: US 60/157,137
/ PRIOR FILING DATE: 1999-09-29
/ PRIOR APPLICATION NUMBER: US 60/163,280
/ PRIOR FILING DATE: 1999-11-03
/ NUMBER OF SEQ ID NOS: 8564
/ SOFTWARE: PatentIn Ver. 3.0
/ SEQ ID NO 5176
/ LENGTH: 128
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-106-698-5176

Query Match          93.1%; Score 27; DB 14; Length 128;
Best Local Similarity 83.3%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GHXXKL 6
Db 119 GHXXKL 124

RESULT 8
US-10-369-493-7208
/ Sequence 7208, Application US/10369493
/ Publication No. US20030233675A1
/ GENERAL INFORMATION:
/ APPLICANT: Cao, Yongwei
/ APPLICANT: Hinkle, Gregory J.
/ APPLICANT: Slater, Steven C.
/ APPLICANT: Goldman, Barry S.
/ APPLICANT: Chen, Xianfeng
/ TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
/ TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
/ FILE REFERENCE: 38-10(52052)B
/ CURRENT APPLICATION NUMBER: US/10/369,493
/ CURRENT FILING DATE: 2003-02-28
/ PRIOR APPLICATION NUMBER: US 60/360,039
/ PRIOR FILING DATE: 2002-02-21
/ NUMBER OF SEQ ID NOS: 47374
/ SEQ ID NO 7208
/ LENGTH: 362
/ TYPE: PRT
/ ORGANISM: Burkholderia cepacia
US-10-369-493-7208
```

Query Match 93.1%; Score 27; DB 15; Length 362;
Best Local Similarity 83.3%; Pred. No. 3.9e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GHXXKL 6
||| ||
Db 152 GHXGKL 157

RESULT 9
US-10-162-335-76
; Sequence 76, Application US/10162335
; Publication No. US20040009480A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, David W. C.
; APPLICANT: Baumgartner, Jason C.
; APPLICANT: Boldog, Ferenc L.
; APPLICANT: Casman, Stacie J.
; APPLICANT: Edinger, Shalom R.
; APPLICANT: Gangolli, Bsha A.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Gorman, Linda
; APPLICANT: Guo, Xiaojia (Sasha)
; APPLICANT: Hjalt, Tord
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Li, Li
; APPLICANT: MacDougall, John R.
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Millet, Isabelle
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Patturajan, Meera
; APPLICANT: Pena, Carol E. A.
; APPLICANT: Rastelli, Luca
; APPLICANT: Shimkets, Richard A.
; APPLICANT: Stone, David J.
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Verniet, Corine A. M.
; APPLICANT: Voss, Edward Z.
; APPLICANT: Zerhusen, Bryan D.
; TITLE OF INVENTION: Therapeutic Polypeptides, Nucleic Acids Encoding Same, and Method
; FILE REFERENCE: 21402-377 B
; CURRENT APPLICATION NUMBER: US/10/162,335
; PRIOR FILING DATE: 2002-10-01
; PRIOR APPLICATION NUMBER: 60/295,507
; PRIOR FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: 60/295,561
; PRIOR FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: 60/296,404
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: 60/296,418
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: 60/297,414
; PRIOR FILING DATE: 2001-06-11
; PRIOR APPLICATION NUMBER: 60/297,567
; PRIOR FILING DATE: 2001-06-12
; PRIOR APPLICATION NUMBER: 60/298,285
; PRIOR FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: 60/298,556
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/299,949
; PRIOR FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: 60/300,883
; PRIOR FILING DATE: 2001-06-26
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 201
; SEQ ID NO 76
; LENGTH: 415
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-162-335-76

Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 GHXXKL 6
||| ||
Db 245 GHXFKL 250

RESULT 10
US-09-801-275-2
; Sequence 2, Application US/09801275
; Patent No. US20020022249A1
; GENERAL INFORMATION:
; APPLICANT: Kapeller-Libermann, Rosana
; TITLE OF INVENTION: 23436. A NOVEL HUMAN UBIQUITIN PROTEASE
; TITLE OF INVENTION: FAMILY MEMBER AND USES THEREOF
; FILE REFERENCE: 10448-021001
; CURRENT APPLICATION NUMBER: US/09/801,275
; CURRENT FILING DATE: 2001-03-06
; PRIOR APPLICATION NUMBER: US 60/187,420
; PRIOR FILING DATE: 2000-03-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 485
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-801-275-2
Query Match 93.1%; Score 27; DB 9; Length 485;
Best Local Similarity 83.3%; Pred. No. 5.2e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 GHXXKL 6
||| ||
Db 306 GHXXKL 311

RESULT 11
US-10-170-789-59
; Sequence 59, Application US/10170789
; Publication No. US20030180930A1
; GENERAL INFORMATION:
; APPLICANT: Rachel E. Meyers
; APPLICANT: Olandt, Peter J.
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Curtis, Rory A. J.
; APPLICANT: Williamson, Mark
; APPLICANT: Welch, Nadine
; TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASE, PHOSPHATASE,
; TITLE OF INVENTION: AND PROTEASE FAMILY MEMBERS AND USES THEREOF
; FILE REFERENCE: 10448-191001
; CURRENT APPLICATION NUMBER: US/10/170,789
; CURRENT FILING DATE: 2002-06-13
; PRIOR APPLICATION NUMBER: US 09/797,039
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: PCT/US01/06525
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/186,061
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 09/882,166
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: PCT/US01/19269
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/212,078
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 09/934,406
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: PCT/US01/26052
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: US 60/226,740
; PRIOR FILING DATE: 2000-08-21
; PRIOR APPLICATION NUMBER: US 09/861,801
; PRIOR FILING DATE: 2001-05-21

; PRIOR APPLICATION NUMBER: PCT/US01/16549
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: US 60/205,508
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: US 09/801,267
; PRIOR FILING DATE: 2001-03-06
; PRIOR APPLICATION NUMBER: PCT/US01/07138
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: US 60/187,454
; PRIOR FILING DATE: 2000-03-07
; PRIOR APPLICATION NUMBER: US 09/829,671
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: PCT/US01/40493
; PRIOR FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: US 60/197,508
; PRIOR FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: US 09/961,721
; PRIOR FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: PCT/US01/29904
; PRIOR FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: US 60/235,023
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US 10/045,367
; PRIOR FILING DATE: 2001-11-07
; PRIOR APPLICATION NUMBER: US 60/246,561
; PRIOR FILING DATE: 2000-11-07
; PRIOR APPLICATION NUMBER: US 09/801,275
; PRIOR FILING DATE: 2001-03-06
; PRIOR APPLICATION NUMBER: PCT/US01/07074
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: US 60/187,420
; PRIOR FILING DATE: 2000-03-07
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 59
; LENGTH: 485
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-170-789-59

Query Match 93.1%; Score 27; DB 14; Length 485;
Best Local Similarity 83.3%; Pred. No. 5.2e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GHKKKL 6
||| ||
Db 306 GHKKKL 311

RESULT 12
US-10-153-668-174
; Sequence 174, Application US/10153668
; Publication No. US20030092616A1
; GENERAL INFORMATION:
; APPLICANT: HONDA, Goichi
; APPLICANT: MATSUDA, Akio
; APPLICANT: MURAMATSU, Shuji
; APPLICANT: ISHIZAWA, Kenya
; TITLE OF INVENTION: STAT6 Activating Gene
; FILE REFERENCE: 1254-0207P
; CURRENT APPLICATION NUMBER: US/10/153,668
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: US 60/293,172
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: US 60/316,031
; PRIOR FILING DATE: 2001-06-31
; PRIOR APPLICATION NUMBER: US 60/328,403
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: JP 2001-157043
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: JP 2001-260681
; PRIOR FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: JP 2001-313175

; PRIOR FILING DATE: 2001-10-10
; NUMBER OF SEQ ID NOS: 488
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 174
; LENGTH: 584
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-153-668-174

Query Match 93.1%; Score 27; DB 14; Length 584;
Best Local Similarity 83.3%; Pred. No. 6.2e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GHKKKL 6
||| ||
Db 370 GHKKKL 375

RESULT 13
US-10-162-335-72
; Sequence 72, Application US/10162335
; Publication No. US20040009480A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, David W.
; APPLICANT: Baumgartner, Jason C.
; APPLICANT: BolDOG, Ferenc L.
; APPLICANT: Casman, Stacie J.
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Gangolli, Bsha A.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Gorman, Linda
; APPLICANT: Guo, Xiaojia (Sasha)
; APPLICANT: HJalt, Tord
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Li, Li
; APPLICANT: MacDougall, John R.
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Millet, Isabelle
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Patturajan, Meera
; APPLICANT: Pena, Carol E. A.
; APPLICANT: Rastelli, Luca
; APPLICANT: Shimkets, Richard A.
; APPLICANT: Stone, David J.
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Vernet, Corine A. M.
; APPLICANT: Voss, Edward Z.
; APPLICANT: Zerhusen, Bryan D.
; TITLE OF INVENTION: Therapeutic Polypeptides, Nucleic Acids Encoding Same, and Methods
; FILE REFERENCE: 21402-377 B
; CURRENT APPLICATION NUMBER: US/10/162,335
; CURRENT FILING DATE: 2002-10-01
; PRIOR APPLICATION NUMBER: 60/295,607
; PRIOR FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: 60/295,661
; PRIOR FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: 60/296,404
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: 60/296,418
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: 60/297,414
; PRIOR FILING DATE: 2001-06-11
; PRIOR APPLICATION NUMBER: 60/297,567
; PRIOR FILING DATE: 2001-06-12
; PRIOR APPLICATION NUMBER: 60/298,285
; PRIOR FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: 60/298,556
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/299,949
; PRIOR FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: 60/300,883
; PRIOR FILING DATE: 2001-06-26
; Remaining Prior Application data removed - See File Wrapper or PALM.

;/ NUMBER OF SEQ ID NOS: 201
;/ SEQ ID NO 72
;/ LENGTH: 615
;/ TYPE: PRT
;/ ORGANISM: Homo sapiens
;/ US-10-162-335-72

Query Match 93.1%; Score 27; DB 15; Length 615;
Best Local Similarity 83.3%; Pred. No. 6.5e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GHXXKL 6
||| ||
Db 445 GHXFKL 450

RESULT 14
US-10-162-335-74
;/ Sequence 74, Application US/10162335
;/ Publication No. US20040009480A1
;/ GENERAL INFORMATION:
;/ APPLICANT: Anderson, David W.
;/ APPLICANT: Baumgartner, Jason C.
;/ APPLICANT: Boldog, Ferenc L.
;/ APPLICANT: Casman, Stacie J.
;/ APPLICANT: Edinger, Shlomit R.
;/ APPLICANT: Gangolli, Esha A.
;/ APPLICANT: Gerlach, Valerie
;/ APPLICANT: Gorman, Linda
;/ APPLICANT: Guo, Xiaojia (Sasha)
;/ APPLICANT: Hjal, Tord
;/ APPLICANT: Kekuda, Ramesh
;/ APPLICANT: Li, Li
;/ APPLICANT: MacDougall, John R.
;/ APPLICANT: Malyaukar, Uriel M.
;/ APPLICANT: Millet, Isabelle
;/ APPLICANT: Padigaru, Muralidhara
;/ APPLICANT: Patturajan, Meera
;/ APPLICANT: Pena, Carol E. A.
;/ APPLICANT: Rastelli, Luca
;/ APPLICANT: Shimkets, Richard A.
;/ APPLICANT: Stone, David J.
;/ APPLICANT: Spytek, Kimberly A.
;/ APPLICANT: Vernet, Corine A. M.
;/ APPLICANT: Voss, Edward Z.
;/ APPLICANT: Zerhusen, Bryan D.
;/ TITLE OF INVENTION: Therapeutic Polypeptides, Nucleic Acids Encoding Same, and Method:
;/ FILE REFERENCE: 21402-377 B
;/ CURRENT APPLICATION NUMBER: US/10/162,335
;/ CURRENT FILING DATE: 2002-10-01
;/ PRIOR APPLICATION NUMBER: 60/295,607
;/ PRIOR FILING DATE: 2001-06-04
;/ PRIOR APPLICATION NUMBER: 60/295,661
;/ PRIOR FILING DATE: 2001-06-04
;/ PRIOR APPLICATION NUMBER: 60/296,404
;/ PRIOR FILING DATE: 2001-06-06
;/ PRIOR APPLICATION NUMBER: 60/296,418
;/ PRIOR FILING DATE: 2001-06-06
;/ PRIOR APPLICATION NUMBER: 60/297,414
;/ PRIOR FILING DATE: 2001-06-11
;/ PRIOR APPLICATION NUMBER: 60/297,567
;/ PRIOR FILING DATE: 2001-06-12
;/ PRIOR APPLICATION NUMBER: 60/298,285
;/ PRIOR FILING DATE: 2001-06-14
;/ PRIOR APPLICATION NUMBER: 60/298,556
;/ PRIOR FILING DATE: 2001-06-15
;/ PRIOR APPLICATION NUMBER: 60/299,949
;/ PRIOR FILING DATE: 2001-06-21
;/ PRIOR APPLICATION NUMBER: 60/300,883
;/ PRIOR FILING DATE: 2001-06-26
;/ Remaining Prior Application data removed - See File Wrapper or PALM.
;/ NUMBER OF SEQ ID NOS: 201
;/ SEQ ID NO 74

;/ LENGTH: 644
;/ TYPE: PRT
;/ ORGANISM: Homo sapiens
;/ US-10-162-335-74

Query Match 93.1%; Score 27; DB 15; Length 644;
Best Local Similarity 83.3%; Pred. No. 6.8e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GHXXKL 6
||| ||
Db 474 GHXFKL 479

RESULT 15
US-10-162-335-84
;/ Sequence 84, Application US/10162335
;/ Publication No. US20040009480A1
;/ GENERAL INFORMATION:
;/ APPLICANT: Anderson, David W.
;/ APPLICANT: Baumgartner, Jason C.
;/ APPLICANT: Boldog, Ferenc L.
;/ APPLICANT: Casman, Stacie J.
;/ APPLICANT: Edinger, Shlomit R.
;/ APPLICANT: Gangolli, Esha A.
;/ APPLICANT: Gerlach, Valerie
;/ APPLICANT: Gorman, Linda
;/ APPLICANT: Guo, Xiaojia (Sasha)
;/ APPLICANT: Hjal, Tord
;/ APPLICANT: Kekuda, Ramesh
;/ APPLICANT: Li, Li
;/ APPLICANT: MacDougall, John R.
;/ APPLICANT: Malyaukar, Uriel M.
;/ APPLICANT: Millet, Isabelle
;/ APPLICANT: Padigaru, Muralidhara
;/ APPLICANT: Patturajan, Meera
;/ APPLICANT: Pena, Carol E. A.
;/ APPLICANT: Rastelli, Luca
;/ APPLICANT: Shimkets, Richard A.
;/ APPLICANT: Stone, David J.
;/ APPLICANT: Spytek, Kimberly A.
;/ APPLICANT: Vernet, Corine A. M.
;/ APPLICANT: Voss, Edward Z.
;/ APPLICANT: Zerhusen, Bryan D.
;/ TITLE OF INVENTION: Therapeutic Polypeptides, Nucleic Acids Encoding Same, and Method:
;/ FILE REFERENCE: 21402-377 B
;/ CURRENT APPLICATION NUMBER: US/10/162,335
;/ CURRENT FILING DATE: 2002-10-01
;/ PRIOR APPLICATION NUMBER: 60/295,607
;/ PRIOR FILING DATE: 2001-06-04
;/ PRIOR APPLICATION NUMBER: 60/295,661
;/ PRIOR FILING DATE: 2001-06-04
;/ PRIOR APPLICATION NUMBER: 60/296,404
;/ PRIOR FILING DATE: 2001-06-06
;/ PRIOR APPLICATION NUMBER: 60/296,418
;/ PRIOR FILING DATE: 2001-06-06
;/ PRIOR APPLICATION NUMBER: 60/297,414
;/ PRIOR FILING DATE: 2001-06-11
;/ PRIOR APPLICATION NUMBER: 60/297,567
;/ PRIOR FILING DATE: 2001-06-12
;/ PRIOR APPLICATION NUMBER: 60/298,285
;/ PRIOR FILING DATE: 2001-06-14
;/ PRIOR APPLICATION NUMBER: 60/298,556
;/ PRIOR FILING DATE: 2001-06-15
;/ PRIOR APPLICATION NUMBER: 60/299,949
;/ PRIOR FILING DATE: 2001-06-21
;/ PRIOR APPLICATION NUMBER: 60/300,883
;/ PRIOR FILING DATE: 2001-06-26
;/ Remaining Prior Application data removed - See File Wrapper or PALM.
;/ NUMBER OF SEQ ID NOS: 201
;/ SEQ ID NO 84
;/ LENGTH: 644
;/ TYPE: PRT

; ORGANISM: Homo sapiens
US-10-162-335-84

Query Match 93.1%; Score 27; DB 15; Length 644;
Best Local Similarity 83.3%; Pred. No. 6.8e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GHKXKL 6
Db 474 GHKFKL 479

Search completed: March 4, 2004, 13:26:33
Job time : 9.33824 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2004, 13:02:33 ; Search time 3.70588 Seconds
(without alignments)
155.739 Million cell updates/sec

Title: SEQ-D
Perfect score: 29
Sequence: 1 ghtxkl 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Maximum Match 0%
Listing first 45 summaries

Database : PIR_78:*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	28	96.6	153	B82309	conserved hypothet
2	28	96.6	218	P84748	probable AP2 domai
3	28	96.6	259	T26298	hypothetical prote
4	28	96.6	304	T02540	probable C2H2-type
5	28	96.6	786	G88065	protein T6A1.2 [i
6	27	93.1	103	G82442	Su11 family protei
7	27	93.1	170	C70541	hypothetical prote
8	27	93.1	177	G70965	hypothetical prote
9	27	93.1	288	H70408	hypothetical prote
10	27	93.1	290	C27115	K-kininogen, LMW p
11	27	93.1	315	A27115	major acute phase
12	27	93.1	375	H63688	response regulator
13	27	93.1	384	B82973	probable peptidic
14	27	93.1	482	B70205	conserved hypothet
15	27	93.1	644	1 KGHU1	kininogen, HMW pre
16	27	93.1	735	T41187	hypothetical prote
17	27	93.1	738	A53542	breifeldin A-sensit
18	27	93.1	840	G83052	translation initia
19	27	93.1	878	F64425	valine-tRNA ligase
20	27	93.1	1096	T21091	hypothetical prote
21	27	93.1	1477	AG3009	polyketide synthet
22	27	93.1	1489	T98274	hypothetical prote
23	27	93.1	1878	B86189	hypothetical prote
24	27	93.1	6713	B89921	hypothetical prote
25	26	89.7	265	T01568	hypothetical prote
26	26	89.7	291	B85018	hypothetical prote
27	26	89.7	380	T01706	hypothetical prote
28	26	89.7	415	C71467	probable tyrosine
29	26	89.7	415	B84544	probable WD-40 rep

ALIGNMENTS

RESULT 1

B82309
conserved hypothetical protein VC0553 [imported] - Vibrio cholerae (strain N16961 serogr
C;Species: Vibrio cholerae
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Nov-2001
C;Accession: B82309
R;Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
Chardson, D.; Ermolaeva, M.B.; Vamathevan, J.; Bass, S.; Qin, H.; Dragci, I.; Sellers, P
1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.

A;Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.

A;Reference number: A82035; MUID:20406833; PMID:10952301

A;Accession: B82309

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-153 <HEI>

A;Cross-references: GB:AE004141; GB:AE003852; NID:g9654976; PIDN:AAF93721.1; GSPDB:GN001:

A;Experimental source: serogroup O1; strain N16961; biotype El Tor

C;Genetics:

A;Gene: VC0553

A;Map position: 1

C;Superfamily: hypothetical protein HI0489

Query Match	96.6%	Score 28;	DB 2;	Length 153;
Best Local Similarity	83.3%	Pred. No. 24;		
Matches	5;	Conservative	0;	Mismatches
		1;	Indels	0;
			Gaps	0;

QY	1	GKXKL	6
Db	89	GKXKL	94

RESULT 2

F84748

probable AP2 domain transcription factor [imported] - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001

C;Accession: F84748

R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; N

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.

euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.

Nature 402, 761-768, 1999

A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A;Reference number: A84420; MUID:20083487; PMID:10617197

A;Accession: F84748

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-218 <STO>

A;Cross-references: GB:AE002093; NID:gl707016; PIDN:AAC69127.1; GSPDB:GN00139

C;Genetics:

A;Gene: At2g3710

A;Map position: 2

Query Match 96.6%; Score 28; DB 2; Length 218;
 Best Local Similarity 83.3%; Pred. No. 34;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GHXXKL 6
 |||||

Db 118 GHXAKL 123
 |||||

RESULT 3

T26298
 hypothetical protein W09C5.1 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 20-Jun-2000
 C:Accession: T26298; T27322
 R:Lennard, N.

submitted to the EMBL Data Library, November 1996
 A:Reference number: Z20192
 A:Accession: T26298
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-259 <WIL>
 A:Cross-references: EMBL:T282077; PIDN:CAB04941.1; GSPDB:GN000019; CESP:W09C5.1
 A:Experimental source: clone W09C5
 R:White, S.
 submitted to the EMBL Data Library, October 1998
 A:Reference number: Z20344
 A:Accession: T27322
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-259 <WIL2>
 A:Cross-references: EMBL:AL032650; PIDN:CRA21705.1; GSPDB:GN000019; CESP:W09C5.1
 A:Experimental source: clone Y635A
 C:Genetics:
 A:Gene: CESP:W09C5.1
 A:Map position: 1
 A:Introns: 24/2; 37/1; 234/3

Query Match 96.6%; Score 28; DB 2; Length 259;
 Best Local Similarity 83.3%; Pred. No. 41;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GHXXKL 6
 |||||

Db 44 GHXAKL 49
 |||||

RESULT 4

T02540
 Probable C2H2-type zinc finger protein At2g37740 [imported] - Arabidopsis thaliana
 N:Alternate names: hypothetical protein F13M22.24
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C>Date: 05-Mar-1999 #sequence_revision 05-Mar-1999 #text_change 16-Feb-2001
 C:Accession: T02540; E94796
 R:Rounsley, S.D.; Kaul, S.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes,
 submitted to the EMBL Data Library, June 1998

A:Description: Arabidopsis thaliana chromosome II BAC F13M22 genomic sequence.
 A:Reference number: Z14677
 A:Accession: T02540
 A>Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-304 <RCU>
 A:Cross-references: EMBL:AC004684; NID:G3236234; PID:G3236256
 A:Experimental source: cultivar Columbia
 R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;
 euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
 Nature 402, 761-768, 1999
 A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
 A:Reference number: A84420; MUID:20083487; PMID:10617197
 A:Accession: E94796
 A>Status: preliminary

A:Molecule type: DNA
 A:Residues: 1-304 <STO>
 A:Cross-references: GB:AE002093; NID:G3236256; PIDN:AAC23644.1; GSPDB:GN00139
 C:Genetics:
 A:Gene: At2g37740; F13M22.24
 A:Map position: 2

Query Match 96.6%; Score 28; DB 2; Length 304;
 Best Local Similarity 83.3%; Pred. No. 48;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GHXXKL 6
 |||||

Db 206 GHXAKL 211
 |||||

RESULT 5

G88065
 protein T16A1.2 [imported] - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
 C:Accession: G88065
 R:Anonymous, The C. elegans Sequencing Consortium.
 Science 282, 2012-2018, 1998

A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biology
 A:Reference number: A75000; MUID:99069613; PMID:9851916
 A:Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/Projects/C.elegans/
 A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
 A:Accession: G88065
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-786 <STO>
 A:Cross-references: GB:chr_II; PIDN:AB37879.1; PID:gl1707126; GSPDB:GN000020; CESP:T16A1.2
 C:Genetics:
 A:Gene: T16A1.2
 A:Map position: 2

Query Match 96.6%; Score 28; DB 2; Length 786;
 Best Local Similarity 83.3%; Pred. No. 1.3e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GHXXKL 6
 |||||

Db 17 GHXTKL 22
 |||||

RESULT 6

G82442
 Sulf family protein VCA0570 [imported] - Vibrio cholerae (strain N16961 serogroup O1)
 C:Species: Vibrio cholerae
 C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
 C:Accession: G82442
 R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
 chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P.
 I., R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
 Nature 406, 477-483, 2000
 A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
 A:Reference number: A82035; MUID:20406833; PMID:10952301

A:Accession: G82442
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-103 <HEI>
 A:Cross-references: GB:AE004388; GB:AE003853; NID:G9657979; PIDN:AAF96472.1; GSPDB:GN001:
 A:Experimental source: serogroup O1; strain N16961; biotype El Tor
 C:Genetics:
 A:Gene: VCA0570
 A:Map position: 2
 C:Superfamily: translation initiation factor SUI

Query Match 93.1%; Score 27; DB 2; Length 103;
 Best Local Similarity 83.3%; Pred. No. 28;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GHKXKL 6
 Db 95 GHKVKL 100

RESULT 7
 C70541
 Hypothetical protein Rv1577c - Mycobacterium tuberculosis (strain H37Rv)
 C:Species: Mycobacterium tuberculosis
 C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
 C:Accession: C70541
 R: Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
 Nature 393, 537-544, 1998
 A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
 A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
 A:Reference number: A70500; MUID:98295987; PMID:9634230
 A:Accession: C70541
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-170 <COL>
 A:Cross-references: GB:Z95586; GB:AL123456; NID:G3261785; PIDN:CAB09060.1; PID:G2117260
 A:Experimental source: strain H37Rv
 C:Genetics:
 A:Gene: Rv1577c

Query Match 93.1%; Score 27; DB 2; Length 170;
 Best Local Similarity 83.3%; Pred. No. 46;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GHKXKL 6
 Db 49 GHKVKL 54

RESULT 8
 G70965
 Hypothetical protein Rv2651c - Mycobacterium tuberculosis (strain H37Rv)
 C:Species: Mycobacterium tuberculosis
 C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
 C:Accession: G70965
 R: Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
 Nature 393, 537-544, 1998
 A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
 A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
 A:Reference number: A70500; MUID:98295987; PMID:9634230
 A:Accession: G70965
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-177 <COL>
 A:Cross-references: GB:Z80225; GB:AL123456; NID:G3242265; PIDN:CAB02356.1; PID:G1550692
 A:Experimental source: strain H37Rv
 C:Genetics:
 A:Gene: Rv2651c

Query Match 93.1%; Score 27; DB 2; Length 177;
 Best Local Similarity 83.3%; Pred. No. 48;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GHKXKL 6
 Db 56 GHKVKL 61

RESULT 9
 H70408
 Hypothetical protein aq_1258 - Aquifex aeolicus
 C:Species: Aquifex aeolicus
 C>Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 05-Nov-1999
 C:Accession: H70408

R: Deckert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Ov V.
 Nature 392, 353-358, 1998
 A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
 A:Reference number: A70300; MUID:98196666; PMID:9537320
 A:Accession: H70408
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-288 <AQF>
 A:Cross-references: GB:AE000731; NID:G2983691; PIDN:AA07261.1; PID:G2983697; GB:AE00065
 A:Experimental source: strain vF5
 C:Genetics:
 A:Gene: aq_1258

Query Match 93.1%; Score 27; DB 2; Length 288;
 Best Local Similarity 83.3%; Pred. No. 78;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GHKXKL 6
 Db 178 GHKVKL 183

RESULT 10
 C27115
 K-kininogen, LMW precursor - rat (fragments)
 C:Species: Rattus norvegicus (Norway rat)
 C>Date: 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change 20-Aug-1999
 C:Accession: C27115; A25488
 R: Fung, W.P.; Schreiber, G.
 J. Biol. Chem. 262, 9298-9308, 1987
 A:Title: Structure and expression of the genes for major acute phase alpha-1-protein (th.
 A:Reference number: A92653; MUID:87250580; PMID:2439509
 A:Accession: C27115
 A:Molecule type: DNA
 A:Residues: 1-290 <FUN>
 R: Kageyama, R.; Kitamura, N.; Ohkubo, H.; Nakanishi, S.
 J. Biol. Chem. 262, 2345-2351, 1987
 A:Title: Differing utilization of homologous transcription initiation sites of rat K and
 A:Reference number: A25488; MUID:87137465; PMID:3818598
 A:Accession: A25488
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-48 <KAG>
 A:Cross-references: GB:J02662; NID:G205071; PIDN:AAA41483.1; PID:G205072
 C:Superfamily: Kininogen; cystatin homology
 F:19-65/Domain: cystatin homology (fragment) <CYS>

Query Match 93.1%; Score 27; DB 2; Length 290;
 Best Local Similarity 83.3%; Pred. No. 79;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GHKXKL 6
 Db 135 GHKVKL 140

RESULT 11
 A27115
 major acute phase alpha-1 protein 1 - rat (fragments)
 C:Species: Rattus norvegicus (Norway rat)
 C>Date: 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change 16-Jul-1999
 C:Accession: A27115
 R: Fung, W.P.; Schreiber, G.
 J. Biol. Chem. 262, 9298-9308, 1987
 A:Title: Structure and expression of the genes for major acute phase alpha-1-protein (th.
 A:Reference number: A92653; MUID:87250580; PMID:2439509
 A:Accession: A27115
 A:Status: not compared with conceptual translation
 A:Molecule type: DNA
 A:Residues: 1-315 <FUN>
 C:Genetics:
 A:Gene: MAP1

C:Superfamily: kininogen; cystatin homology
F:19-65/Domain: cystatin homology (fragment) <Cys>

Query Match 93.1%; Score 27; DB 2; Length 315;
Best Local Similarity 83.3%; Pred. No. 85;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GHKXKL 6
|||
Db 159 GHKXKL 164

RESULT 12

H59688 response regulator aspartate phosphatase rapE - Bacillus subtilis

C:Species: Bacillus subtilis
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 21-Jul-2000
C:Accession: H59688
R:Kunst, P.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berton
C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho
A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallier
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, F.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron
A:Authors: Schleich, S.; Schroeter, R.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
akeuchi, M.; Tanakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.
A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A:Reference number: A69580; MUID:98044033; PMID:9384377
A:Accession: H59688
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-375 <KUN>
A:Cross-references: GB:Z99117; GB:AL009126; NID:g2634966; PIDN:CABI4524.1; PID:g2635028
A:Experimental source: strain 169
C:Genetics:
A:Gene: rapE
C:Superfamily: response regulator aspartate phosphatase rapA

Query Match 93.1%; Score 27; DB 1; Length 375;
Best Local Similarity 83.3%; Pred. No. 1e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GHKXKL 6
|||
Db 85 GHKXKL 90

RESULT 13

B82973 probable peptidic bond hydrolase PA5390 [imported] - Pseudomonas aeruginosa (strain PA01

C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 03-Feb-2003
C:Accession: B82973
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Br
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A:Reference number: A82950; MUID:20437337; PMID:10384043
A:Accession: B82973
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-384 <STO>
A:Cross-references: GB:AE004951; GB:AE004091; NID:g9951705; PIDN:AA608775.1; GSPDB:GN001
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA5390

C:Superfamily: succinyl-diaminopimelate desuccinylase

Query Match 93.1%; Score 27; DB 2; Length 384;
Best Local Similarity 83.3%; Pred. No. 1e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GHKXKL 6
|||
Db 171 GHKXKL 176

RESULT 14

B70205

conserved hypothetical integral membrane protein BB0843 - Lyme disease spirochete

C:Species: Borrelia burgdorferi (Lyme disease spirochete)
C:Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 08-Oct-1999
C:Accession: B70205
R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; White,
son, D.; Peterson, J.; Kierlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt,
; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.
Nature 390, 580-586, 1997
A:Authors: Smith, H.O.; Venter, J.C.
A:Title: Genomic sequence of a Lyme disease spirochete, Borrelia burgdorferi.
A:Reference number: A70100; MUID:98065943; PMID:9403685
A:Accession: B70205
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-482 <KLE>
A:Cross-references: GB:AE000183; GB:AE000783; NID:g2688786; PIDN:AA67189.1; PID:g2688787
A:Experimental source: strain B31
C:Superfamily: Haemophilus influenzae conserved hypothetical protein HI0594

Query Match 93.1%; Score 27; DB 2; Length 482;
Best Local Similarity 83.3%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GHKXKL 6
|||
Db 122 GHKXKL 127

RESULT 15

KGHUHI

kininogen, HMW precursor [validated] - human

N:Alternate names: alpha-2-thiol proteinase inhibitor; preprokinogen; prokinogen
N:Contains: bradykinin (kallidin I); HMW kininogen I; HMW kininogen II; low molecular we
C:Species: Homo sapiens (man)
C:Date: 28-May-1986 #sequence_revision 28-May-1986 #text_change 08-Dec-2000
C:Accession: A01279; A52276; S32422; A91153; A24871; A27699; A31905; A34030; S02
R:Ohkubo, I.; Kurachi, K.; Takasawa, T.; Shiokawa, H.; Sasaki, M.
Biochemistry 23, 5691-5697, 1984
A:Title: Isolation of a human cDNA for alpha-2-thiol proteinase inhibitor and its identifi
A:Reference number: A90490; MUID:85122621; PMID:6441591
A:Accession: A01279
A:Molecule type: mRNA
A:Residues: 1-389 <OHK>
A:Cross-references: GB:X02566; NID:g177889
R:Takagaki, Y.; Kitamura, N.; Nakanishi, S.
J. Biol. Chem. 260, 8601-8609, 1985

A:Title: Cloning and sequence analysis of cDNAs for human high molecular weight and low
A:Reference number: A92544; MUID:85234582; PMID:2989293
A:Accession: A52276
A:Molecule type: mRNA
A:Residues: 1-592, '1', 594-644 <YAK>
A:Cross-references: GB:M11437; NID:g186751; PIDN:AA859550.1; PID:g386852
R:Werswald, E.A.; Roessler, D.; Mentele, R.; Assfalg-Machleidt, I.
FEBS Lett. 321, 93-97, 1993

A:Title: Cloning, expression and characterization of human kininogen domain 3.
A:Reference number: S32422; MUID:9323854; PMID:8467916
A:Accession: S32422
A:Molecule type: mRNA
A:Residues: 'ANSM', 253-377 <AUE>
A:Note: differences are due to known cloning artifacts

R;Lottspeich, F.; Kellermann, J.; Henschen, A.; Foerlsch, B.; Muller-Esterl, W.
 Eur. J. Biochem. 152, 307-314, 1985
 A;Title: The amino acid sequence of the light chain of human high-molecular-mass kininogen
 A;Reference number: A91153; MUID:86030270; PMID:4054110
 A;Accession: A91153
 A;Molecule type: protein
 A;Residues: 379-644 <LOT>
 A;Note: the bradykinin sequence preceding the light chain sequence was not determined in
 R;Kellermann, J.; Lottspeich, F.; Henschen, A.; Mueller-Esterl, W.
 Eur. J. Biochem. 154, 471-478, 1986
 A;Title: Completion of the primary structure of human high-molecular-mass kininogen. The
 A;Reference number: A24871; MUID:86108361; PMID:3484703
 A;Accession: A24871
 A;Molecule type: protein
 A;Residues: 1-19, 20-380 <KEL1>
 R;Kellermann, J.; Lottspeich, F.; Henschen, A.; Mueller-Esterl, W.
 In Kinins IV, Greenbaum, L.M., and Margolius, H.S., ed., pp.85-89, Plenum Press, New York
 A;Title: Amino acid sequence of the light chain of human high molecular mass kininogen.
 A;Reference number: A27899
 A;Accession: A27899
 A;Molecule type: protein
 A;Residues: 379-389, K', 390-407, Q', 409-644 <KEL2>
 R;Mindrou, T.; Carretero, O.A.; Proud, D.; Walz, D.; Scicli, A.G.
 Biochem. Biophys. Res. Commun. 152, 519-526, 1988
 A;Title: A new kinin moiety in human plasma kininogens.
 A;Reference number: A27699; MUID:88209021; PMID:335237
 A;Accession: A27699
 A;Molecule type: protein
 A;Residues: 380-389 <MIN>
 R;Maeda, H.; Matsumura, Y.; Kato, H.
 J. Biol. Chem. 263, 16051-16054, 1988
 A;Title: Purification and identification of [hydroxyprolyl(3)]bradykinin in ascitic fluid
 A;Reference number: A31905; MUID:89034061; PMID:3132782
 A;Accession: A31905
 A;Molecule type: protein
 A;Residues: 381-389 <MAE>
 R;Sasaguri, M.; Ikeda, M.; Ideishi, M.; Arakawa, K.
 Biochem. Biophys. Res. Commun. 150, 511-516, 1988
 A;Title: Identification of [hydroxyproline(3)]-lysyl-bradykinin released from human plas
 A;Reference number: A34030; MUID:88106632; PMID:3337729
 A;Accession: A34030
 A;Molecule type: protein
 A;Residues: 380-389 <SAS>
 R;Lenarcic, B.; Gabrijelcic, D.; Rozman, B.; Drobnic-Kosorok, M.; Turk, V.
 Biol. Chem. Hoppe-Seyler 369, 257-261, 1988
 A;Title: Human cathepsin B and cysteine proteinase inhibitors (CPIs) in inflammatory and
 A;Reference number: S02482; MUID:89076517; PMID:3284507
 A;Accession: S02482
 A;Molecule type: protein
 A;Residues: 1-19, 189-192, 310-314, 381-389 <LEN1>
 R;Kato, H.; Matsumura, Y.; Maeda, H.
 FEBS Lett. 232, 252-254, 1988
 A;Title: Isolation and identification of hydroxyproline analogues of bradykinin in human
 A;Reference number: A61495; MUID:88211869; PMID:3366244
 A;Accession: A61495
 A;Molecule type: protein
 A;Residues: 380-389 <KAT1>
 A;Experimental source: urine
 A;Note: this peptide had Pro-383 modified to 4-hydroxyproline
 A;Accession: B61495
 A;Molecule type: protein
 A;Residues: 381-389 <KAT2>
 A;Experimental source: urine
 A;Note: this peptide had Pro-383 modified to 4-hydroxyproline
 A;Accession: C61495
 A;Molecule type: protein
 A;Residues: 380-389 <KAT3>
 R;Lenarcic, B.; Krasovec, M.; Ritonja, A.; Olafsson, I.; Turk, V.
 FEBS Lett. 280, 211-215, 1991
 A;Title: Inactivation of human cystatin C and kininogen by human cathepsin D.
 A;Reference number: S14303; MUID:91192133; PMID:2013314
 A;Accession: S14303
 A;Molecule type: protein

A;Residues: 264-359, N', 361-375 <LEN2>
 R;Little, S.S.; Johnson, D.A.
 Biochem. J. 307, 341-346, 1995
 A;Title: Human mast cell tryptase isoforms: separation and examination of substrate-spec
 A;Reference number: S55239; MUID:95251593; PMID:7733867
 A;Accession: S55239
 A;Molecule type: protein
 A;Residues: 450-452, X', 454, X', 456 <LIT>
 R;Straczek, J.; Maachi, F.; Nguyen, D.; Becchi, M.; Heulin, M.H.; Nabet, P.; Bellevil
 FEBS Lett. 373, 207-211, 1995
 A;Title: Purification from human plasma of a tetrapeptide that potentiates insulin-like
 A;Reference number: S68059; MUID:96033974; PMID:7589467
 A;Accession: S68059
 A;Molecule type: protein
 A;Residues: 431-434 <STR>
 R;Kitamura, N.; Kitagawa, H.; Fukushima, D.; Takagaki, Y.; Miyata, T.; Nakanishi, S.
 J. Biol. Chem. 260, 8610-8617, 1985
 A;Title: Structural organization of the human kininogen gene and a model for its evoluti
 A;Reference number: A92545; MUID:85234583; PMID:2989294
 A;Contents: annotation; gene organization
 R;Pierce, J.V.
 Fed. Proc. 27, 52-57, 1968
 A;Title: Structural features of plasma kinins and kininogens.
 A;Reference number: A91455; MUID:90255622; PMID:4952632
 A;Contents: annotation; bradykinin
 C;Comment: The HMW kininogen precursor and the LMW form are produced from the same gene
 C;Comment: Kininogen is a cysteine proteinase inhibitor, takes part in initiation of the
 C;Comment: The glycine/histidine/lysine-rich region of HMW kininogen light chain is impo
 C;Comment: Bradykinin, released from kininogen by kallikrein, is a potent vasodilator, 1
 xyproline residue is present in the kininogen prior to the release of bradykinin.
 C;Genetics:
 A;Gene: GDB:KNG
 A;Cross-references: GDB:125256; OMIM:228960
 A;Map Position: 3q27-3q27
 A;Intons: 65/3; 102/3; 131/1; 188/3; 224/3; 253/1; 310/3; 346/3; 375/3
 C;Superfamily: kininogen; cystatin homology
 C;Keywords: alternative splicing; blood coagulation; cysteine proteinase inhibitor; dupl
 F;1-18/Domain: signal sequence #status experimental <SIG>
 F;19-644/Product: HMW kininogen I (prokininogen) #status experimental <MAT1>
 F;19-379,390-644/Product: HMW kininogen II #status experimental <MAT2>
 F;19-379/Domain: HMW kininogen heavy chain #status experimental <HCH>
 F;19-131/Domain: cystatin homology <CY1>
 F;142-253/Domain: cystatin homology <CY2>
 F;264-375/Domain: cystatin homology <CY3>
 F;380-389/Product: lysyl-bradykinin (kallidin I) #status experimental <BDY>
 F;381-389/Product: bradykinin (kallidin I) #status experimental <BDY>
 F;390-644/Domain: HMW kininogen light chain #status experimental <LCH>
 F;421-510/Region: glycine/histidine/lysine-rich 30-residue repeats
 F;431-434/Product: low molecular weight growth promoting factor #status experimental <GP
 F;19/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experimen
 F;28-614, 83-94, 107-126, 142-145, 206-218, 229-248, 264-267, 328-340, 351-370/Disulfide bonds: :
 F;48/Binding site: carbohydrate (Asn) (covalent) #status absent
 F;169,205,294/Binding site: carbohydrate (Asn) (covalent) #status experimental
 F;379-380/Cleavage site: Met-Lys (kallikrein) #status experimental
 F;383/Modified site: 4-hydroxyproline (Pro) (partial) #status experimental
 F;389-390/Cleavage site: Arg-Ser (kallikrein) #status experimental
 F;401,533,542,546,557,571,593,628/Binding site: carbohydrate (Thr) (covalent) #status ex
 F;577/Binding site: carbohydrate (Ser) (covalent) #status experimental

Query Match 93.1%; Score 27; DB 1; Length 644;
 Best Local Similarity 83.3%; Pred. No. 1.8e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GHKXKL 6
 DB 474 GHKFL 479

Search completed: March 4, 2004, 13:11:24
 Job time : 4.70588 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2004, 12:53:12 ; Search time 2.38235 Seconds
(without alignments)
131.140 Million cell updates/sec

Title: SEQ-D
Perfect score: 29
Sequence: 1 ghkxkl 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	28	96.6	232	1	CMPT CITNA
2	27	93.1	375	1	RAPE BACSU
3	27	93.1	644	1	KNG HUMAN
4	27	93.1	735	1	YQ37 SCHPO
5	27	93.1	738	1	COG2 HUMAN
6	27	93.1	840	1	IF2 PSEAE
7	27	93.1	841	1	IF2 WIGBR
8	27	93.1	878	1	SIV METUA
9	26	89.7	415	1	MS12 ARATH
10	26	89.7	424	1	MS13 ARATH
11	26	89.7	445	1	YAV5 XANCV
12	26	89.7	803	1	IF2 COXBU
13	26	89.7	2452	1	RPE1 PIAPD
14	25	86.2	154	1	RL22 HALMA
15	25	86.2	262	1	VIRF SHIFL
16	25	86.2	371	1	THI1 SULSO
17	25	86.2	436	1	SEPT RAT
18	25	86.2	451	1	GSHR PSEAE
19	25	86.2	457	1	V5IK ACLSA
20	25	86.2	543	1	IP09 YEAST
21	25	86.2	556	1	YEXO YEAST
22	25	86.2	583	1	CRTI LYCRS
23	25	86.2	619	1	DXS CLOPE
24	25	86.2	700	1	NCD DROME
25	25	86.2	841	1	IF2 PSESM
26	25	86.2	861	1	PQ58 CAEEL
27	25	86.2	972	1	TOPI DROME
28	25	86.2	1123	1	V120 HSV11
29	25	86.2	2332	1	POLG FMDVA
30	25	86.2	4289	1	TENX HUMAN
31	24	82.8	121	1	RBFA CLOTE
32	24	82.8	123	1	VMAT VHSVO
33	24	82.8	124	1	YGBA_ECOLI

34	24	82.8	158	1	Y236 MYCPN
35	24	82.8	171	1	IF3 TETH
36	24	82.8	196	1	PTH VIBCH
37	24	82.8	196	1	PTH VIBPA
38	24	82.8	196	1	PTH VIBVU
39	24	82.8	200	1	CYDX HUMAN
40	24	82.8	200	1	CYDX RAT
41	24	82.8	210	1	FLPA ARCFU
42	24	82.8	211	1	YB48 METAC
43	24	82.8	211	1	YP34 METWA
44	24	82.8	212	1	IF3 DEIRA
45	24	82.8	236	1	KDGM ERWCH

ALIGNMENTS

RESULT 1
CAMT CITNA
ID CAMT CITNA STANDARD; PRT; 232 AA.
AC Q9SLP8;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Caffeoyl-CoA O-methyltransferase (EC 2.1.1.104) (Trans-caffeoyl-CoA
3-O-methyltransferase) (CCoAMT) (CCoAMT)
OS Citrus natsudaidai (Natsudaidai orange) (Japanese summer orange).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eucosids II; Sapindales; Rutaceae; Citrus.
OX NCBI_TaxID=109792;
RN [1]
RP SEQUENCE FROM N.A.
RA Gryciuk A.A., Tsuyumu S.;
RT "Cloning of a genomic DNA encoding caffeoyl-CoenzymeA 3-O-
methyltransferase of citrus.";
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Methylates caffeoyl-CoA to feruloyl-CoA and 5-
hydroxyferuloyl-CoA to sinapoyl-CoA. Plays a role in the synthesis
of feruloylated polysaccharides. Involved in the reinforcement of
the plant cell wall, also involved in the responding to wounding
or pathogen challenge by the increased formation of cell wall-
bound ferulic acid polymers.
CC -!- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + caffeoyl-CoA = S-
adenosyl-L-homocysteine + feruloyl-CoA.
CC -!- PATHWAY: Lignin biosynthesis.
CC -!- PATHWAY: Phenylpropanoid pathway.
CC -!- SIMILARITY: Belongs to the methyltransferase superfamily. Family
3. CCoAMT subfamily.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; AB035144; BAA88234.1; -;
InterPro; IPR002935; Methyltransf_3.
InterPro; IPR000051; SAM_bind.
Pfam; PF01596; Methyltransf_3; 1.
Lignin biosynthesis; Transferase; Methyltransferase.
SQ SEQUENCE 232 AA; 26333 MW; DAB002D0CA1101EF CRC64;

Query Match 96.6%; Score 28; DB 1; Length 232;
Best Local Similarity 83.3%; Pred. No. 14;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 GHKXKL 6

Db 4 GHKSKL 9


```

RESULT 2
RAPE BACSU STANDARD; PRT; 375 AA.
AC P45943;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Response regulator aspartate phosphatase E (EC 3.1.-.-).
GN RAPE OR BSU25830.
OS Bacillus subtilis.
OX Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
ON NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168 / JH642;
RX MEDLINE=95219086; PubMed=7704261;
RA Takemaru K.-I., Mizuno M., Sato T., Takeuchi M., Kobayashi Y.;
RT "Complete nucleotide sequence of a skin element excised by DNA
RT rearrangement during sporulation in Bacillus subtilis.";
RL Microbiology 141:323-327(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=168 / JH642;
RX MEDLINE=97124195; PubMed=8969508;
RA Mizuno M., Masuda S., Takemaru K.-I., Hosono S., Sato T., Takeuchi M.,
RA Kobayashi Y.;
RT Systematic sequencing of the 283 kb 210 degrees-232 degrees region of
RT the Bacillus subtilis genome containing the skin element and many
RT sporulation genes.";
RL Microbiology 142:3103-3111(1996).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borcherdt S.,
RA Borriess R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Brusch C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Conerton I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,
RA Entiaz K., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz S.V., Glaser P., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Gim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
RA Guisepi G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Mauel J., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogawa A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA Prescan E., Fujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA Sekiguchi J., Sekowska A., Seror S.J., Serron P., Shin B.S., Soldo B.,
RA Sorokin A., Taccon E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tanakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.;
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
RT subtilis.";
RL Nature 390:249-256(1997).
RN [4]
RP IDENTIFICATION.
RX MEDLINE=96084975; PubMed=7489895;
RA Medigue C., Moszer I., Viari A., Danchin A.;
RT "Analysis of a Bacillus subtilis genome fragment using a co-operative
RT computer system prototype.";
RL Gene 165:GC37-GC51(1995).

```

```

CC -1- SIMILARITY: Belongs to the RAP family.
CC -1- SIMILARITY: Contains 6 TPR repeats.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D32216; BAA06965.1; -.
DR EMBL; D84432; BAA12429.1; -.
DR EMBL; Z99117; CAB14524.1; -.
DR PIR; H69688; H69688.
DR Subtilisin; BG11299; rapE.
DR InterPro; IPR008940; Prenyl_trans.
DR InterPro; IPR001440; TPR.
DR Pfam; PF00515; TPR; 4.
DR SMART; SM00028; TPR; 3.
KW Hydrolase; Repeat; TPR repeat; Complete proteome.
FT REPEAT 96 129
FT REPEAT 177 210
FT REPEAT 219 252
FT REPEAT 258 291
FT REPEAT 297 330
FT REPEAT 333 366
FT REPEAT 375 44564 MW; 9AD9D33C012DFDA8 CRC64;
SQ SEQUENCE 375 AA; 44564 MW; 9AD9D33C012DFDA8 CRC64;
Query Match 93.1%; Score 27; DB 1; Length 375;
Best Local Similarity 83.3%; Pred. No. 40;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GHXXKL 6
DB 85 GHXXKL 90
KING_HUMAN STANDARD; PRT; 644 AA.
ID KING_HUMAN STANDARD; PRT; 644 AA.
AC P01042; P01043;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Kininogen precursor (alpha-2-thiol proteinase inhibitor) [Contains:
DE Bradykinin].
GN KNG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS HMW AND LMW).
RC TISSUE=Liver;
RX MEDLINE=85234582; PubMed=2989293;
RA Takagaki Y., Kitamura N., Nakanishi S.;
RT "Cloning and sequence analysis of cDNAs for human high molecular
RT weight and low molecular weight prekininogens. Primary structures of
RT two human prekininogens.";
RL J. Biol. Chem. 260:8601-8609(1985).
RN [2]
RP GENE STRUCTURE.
RX MEDLINE=85234583; PubMed=2989294;
RA Kitamura N., Kitagawa H., Fukushima D., Takagaki Y., Miyata T.,
RA Nakanishi S.;
RT "Structural organization of the human kininogen gene and a model for
RT its evolution.";
RL J. Biol. Chem. 260:8610-8617(1985).
RN [3]
RP SEQUENCE OF 1-401 FROM N.A.
RX MEDLINE=8512621; PubMed=6441591;
RA Chkubo I., Kurachi K., Takasawa T., Shiokawa H., Sasaki M.;

```

RT "Isolation of a human cDNA for alpha 2-thiol proteinase inhibitor and
 its identity with low molecular weight kininogen.";
 RL Biochemistry 23:5691-5697(1984).
 RN [4]
 RP SEQUENCE OF 379-644.
 RX MEDLINE=86030270; PubMed=4054110;
 RA Lottspeich F., Kellermann J., Henschen A., Foerster B.,
 Mueller-Esterl W.;
 RT "The amino acid sequence of the light chain of human high-molecular-
 mass kininogen.";
 RL Eur. J. Biochem. 152:307-314(1985).
 RN [5]
 RP SEQUENCE OF 381-389.
 RX MEDLINE=90255622; PubMed=4952632;
 RA Pierce J.V.;
 RT "Structural features of plasma kinins and kininogens.";
 RL Fed. Proc. 27:52-57(1968).
 RN [6]
 RP DISULFIDE BONDS.
 RA Sueyoshi T., Miyata T., Kato H., Iwanaga S.;
 RT "Disulfide bonds in bovine HMW kininogens.";
 RL Seikagaku 56:808-808(1984).
 RN [7]
 RP CARBOHYDRATE-LINKAGE SITE ASN-294.
 RX MEDLINE=22660472; PubMed=12754519;
 RA Zhang H., Li X.-J., Martin D.B., Aebersold R.;
 RT "Identification and quantification of N-linked glycoproteins using
 hydrazide chemistry, stable isotope labeling and mass spectrometry.";
 RL Nat. Biotechnol. 21:660-666(2003).
 CC -!- FUNCTION: (1) Kininogens are inhibitors of thiol proteases; (2)
 CC HMW-kininogen plays an important role in blood coagulation by
 CC helping to position optimally prekallikrein and factor XI next to
 CC factor XII; (3) HMW-kininogen inhibits the thrombin- and plasmin-
 CC induced aggregation of thrombocytes; (4) the active peptide
 CC bradykinin that is released from HMW-kininogen shows a variety of
 CC physiological effects: (4A) influence in smooth muscle
 CC contraction, (4B) induction of hypotension, (4C) natriuresis and
 CC diuresis, (4D) decrease in blood glucose level, (4E) it is a
 CC mediator of inflammation and causes (4E1) increase in vascular
 CC permeability, (4E2) stimulation of nociceptors (4E3) release of
 CC other mediators of inflammation (e.g. prostaglandins), (4F) it has
 CC a cardioprotective effect (directly via bradykinin action,
 CC indirectly via endothelium-derived relaxing factor action); (5)
 CC HMW-kininogen inhibits the aggregation of thrombocytes; (6) LMW-
 CC kininogen is in contrast to HMW-kininogen not involved in blood
 CC clotting.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=HMW;
 CC IsoId=P01042-1; Sequence=Displayed;
 CC Name=LMW;
 CC IsoId=P01042-2; Sequence=VSP_001261; VSP_001262;
 CC -!- TISSUE SPECIFICITY: Plasma.
 CC -!- PTM: Bradykinin is released from kininogen by plasma kallikrein.
 CC -!- SIMILARITY: Contains 3 cystatin-like domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; K02566; AAB59550.1; JOINED.
 CC EMBL; M11437; AAB59550.1; JOINED.
 CC EMBL; M11438; AAB59550.1; JOINED.
 CC EMBL; M11521; AAB59550.1; JOINED.
 CC EMBL; M11522; AAB59550.1; JOINED.
 CC EMBL; M11523; AAB59550.1; JOINED.
 CC EMBL; M11524; AAB59550.1; JOINED.
 CC EMBL; M11525; AAB59550.1; JOINED.
 CC EMBL; M11526; AAB59550.1; JOINED.

DR EMBL; M11526; AAB59550.1; JOINED.
 DR EMBL; M11527; AAB59550.1; JOINED.
 DR EMBL; M11528; AAB59550.1; JOINED.
 DR EMBL; M11437; AAB59550.1; JOINED.
 DR EMBL; M11438; AAB59550.1; JOINED.
 DR EMBL; M11521; AAB59550.1; JOINED.
 DR EMBL; M11522; AAB59550.1; JOINED.
 DR EMBL; M11523; AAB59550.1; JOINED.
 DR EMBL; M11524; AAB59550.1; JOINED.
 DR EMBL; M11525; AAB59550.1; JOINED.
 DR EMBL; M11526; AAB59550.1; JOINED.
 DR EMBL; M11527; AAB59550.1; JOINED.
 DR EMBL; M11528; AAB59550.1; JOINED.
 DR PIR; A01279; KGEUHI.
 DR PIR; A01280; KGEUHI.
 DR SWISS-2DPAGE; P01042; HUMAN.
 DR Genew; HGNC:6383; KNG.
 DR MM; 228960; -.
 DR GO; GO:0007596; P: blood coagulation; NAS.
 DR GO; GO:0030146; P: diuresis; NAS.
 DR GO; GO:0006954; P: inflammatory response; NAS.
 DR GO; GO:0030147; P: natriuresis; NAS.
 DR GO; GO:0006939; P: smooth muscle contraction; NAS.
 DR InterPro; IPR000010; Cystatin.
 DR InterPro; IPR002395; Kininogen.
 DR Pfam; PF00031; Cystatin; 3.
 DR PRINTS; PR00334; KININOGEN.
 DR SMART; SM00043; CY; 3.
 DR PROSITE; PS00287; CYSTATIN; 2.
 KW Glycoprotein; Plasma; Repeat; Thiol protease inhibitor; Vasodilator;
 KW Bradykinin; Blood coagulation; Inflammatory response; Signal;
 KW Alternative splicing; Pyrrolidone carboxylic acid.
 FT SIGNAL 1 18
 FT CHAIN 19 644 KININOGEN
 FT CHAIN 19 380 KININOGEN HEAVY CHAIN.
 FT PEPTIDE 381 389 BRADIKININ.
 FT CHAIN 390 644 KININOGEN LIGHT CHAIN.
 FT DOMAIN 19 136 CYSTATIN-LIKE 1.
 FT DOMAIN 137 258 CYSTATIN-LIKE 2.
 FT DOMAIN 259 380 CYSTATIN-LIKE 3.
 FT DOMAIN 420 510 HIS-RICH
 FT REPEAT 420 449 (ASSOCIATED WITH CLOTTING ACTIVITY).
 FT REPEAT 450 479
 FT REPEAT 480 510
 FT MOD RES 19 19
 FT DISULFID 28 614
 FT DISULFID 83 94
 FT DISULFID 107 126
 FT DISULFID 142 145
 FT DISULFID 206 218
 FT DISULFID 229 248
 FT DISULFID 264 267
 FT DISULFID 328 340
 FT DISULFID 351 370
 FT CARBOHYD 48 48
 FT CARBOHYD 169 169
 FT CARBOHYD 205 205
 FT CARBOHYD 294 294
 FT CARBOHYD 401 401
 FT CARBOHYD 533 533
 FT CARBOHYD 542 542
 FT CARBOHYD 546 546
 FT CARBOHYD 557 557
 FT CARBOHYD 571 571
 FT CARBOHYD 577 577
 FT CARBOHYD 593 593
 FT CARBOHYD 628 628
 FT VARSPLIC 402 427
 FT VARSPLIC 428 644
 FT Missing (in isoform LMW).
 FT /FTId=VSP_001261
 FT /FTId=VSP_001262.

VSPPTSMAQDERDSGKEQGHTR -> SHLRSCYKGR
 PKKAGAEPAEREVS (in isoform LMW).
 Missing (in isoform LMW).
 /FTId=VSP_001262.

FT CONFLICT 593 593 T -> I (IN REF. 1).
SQ SEQUENCE 644 AA; 71945 MW; 3132B4CB4F8PB7E CRC64;
Query Match 93.1%; Score 27; DB 1; Length 644;
Best Local Similarity 83.3%; Pred. No. 72;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GHKXKL 6
DB 474 GHKFKL 479
RESULT 4
ID YQ37 SCHPO STANDARD; PRT; 735 AA.
AC Q10432;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein C188.07 in chromosome III.
GN SPCC188.07.
OS Schizosaccharomyces pombe (Fission yeast).
OC Bkaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajadream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouras J., Hayes J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Rucke E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds R., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
RA Welljens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Cabell C., Fuchs M., Fritz C., Holzer E., Moest L., Hilbert H.,
RA Borzym K., Langer I., Beck A., Leirach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Fumelle B.,
RA Geffreau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gallard C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., Gel Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe."
RL Nature 415:871-880(2002).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AL049662; CAB41227.1; -.
DR PIR; T41187; T41187.
DR GeneDB SPombe; SPCC188.07; -.
KW Hypothetical protein.
SQ SEQUENCE 735 AA; 83300 MW; 14193682700A2634 CRC64;
Query Match 93.1%; Score 27; DB 1; Length 735;
Best Local Similarity 83.3%; Pred. No. 83;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GHKXKL 6
DB 52 GHKFKL 57
RESULT 5
ID COG2 HUMAN STANDARD; PRT; 738 AA.
AC Q14746;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Conserved oligomeric Golgi complex component 2 (Low density
DE lipoprotein receptor defect C-complementing protein).
GN COG2 OR LDLC.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95050941; PubMed=7962052;
RA Podos S.D., Reddy P., Astkenas J., Krieger M.;
RT "LDLC encodes a brefeldin A-sensitive, peripheral Golgi protein
RT required for normal Golgi function."
RL J. Cell Biol. 127:679-691(1994).
RN [2]
RP SEQUENCE FROM N.A.
RX TISSUE=Retinal pigment epithelium;
RC MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max J., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Millaby S.J.,
RA Bosak S.A., McWeary P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S.I., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Munz D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Maizumi M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -1- FUNCTION: Required for normal Golgi morphology and function.
CC -1- SUBUNIT: Component of the conserved oligomeric Golgi complex which
CC is composed of eight different subunits and is required for normal
CC Golgi morphology and localization.
CC -1- SUBCELLULAR LOCATION: Associated with the cytosolic face of
CC the Golgi apparatus (Probable).
CC -1- SIMILARITY: Belongs to the COG2 family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z34975; CAB4427.1; -.
DR EMBL; BC014960; AAE14960.1; -.
DR PIR; A33542; A33542.
DR Genew; HGNC:6546; COG2.
DR MIM; 606974; -.

DR GO; GO:0000139; C:Golgi membrane; TAS.
 DR GO; GO:0017119; C:Golgi transport complex; IEP.
 DR GO; GO:0008565; P:protein transporter activity; IMP.
 DR GO; GO:0007030; P:Golgi organization and biogenesis; IMP.
 DR GO; GO:0006891; P:intra-Golgi transport; IMP.
 DR GO; GO:0006886; P:intracellular protein transport; TAS.
 DR GO; GO:0009312; P:oligosaccharide biosynthesis; TAS.
 DR GO; GO:0006486; P:protein amino acid glycosylation; TAS.
 KW Transport; Protein transport; Golgi stack; Membrane.
 SQ SEQUENCE 738 AA; 83207 MW; 8E393CBF1114DB28 CRC64;

Query Match 93.1%; Score 27; DB 1; Length 738;
 Best Local Similarity 83.3%; Pred. No. 84;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GHKXKL 6
 |||||
 Db 621 GHKDL 626

RESULT 6
 IF2_PSEAE
 ID IF2_PSEAE STANDARD; PRT; 840 AA.
 AC QSHV55;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Translation initiation factor IF-2.
 GN INFB OR PA4744.
 OS Pseudomonas aeruginosa.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Pseudomonas.
 OX NCBI_TaxID=287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 15692 / PA01;
 RX MEDLINE=20437337; PubMed=10984043;
 RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,
 RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
 RA Garber R.L., Goltzy L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
 RA Brody L.L., Goltzy S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
 RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
 RT "complete genome sequence of Pseudomonas aeruginosa PA01, an
 RT opportunistic pathogen.";
 RL Nature 406:959-964(2000).
 CC -!- FUNCTION: One of the essential components for the initiation of
 CC protein synthesis. Protects formylmethionyl-tRNA from spontaneous
 CC hydrolysis and promotes its binding to the 30S ribosomal subunits.
 CC Also involved in the hydrolysis of GTP during the formation of the
 CC 70S ribosomal complex (By similarity).
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -!- SIMILARITY: Belongs to the IF-2 family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR ENBL; AE004888; AAG08130.1; -.
 DR HAMAP; G83052; G83052.
 DR PIR; G83052; -. 1.
 DR InterPro; IPR000795; EF_GTPbind.
 DR InterPro; IPR004161; EFTU_D2.
 DR InterPro; IPR000178; IF2.
 DR InterPro; IPR006847; IF2 N.
 DR InterPro; IPR005225; Small GTP.
 DR Pfam; PF00009; GTP_EFTU; 1.
 DR Pfam; PF03144; GTP_EFTU_D2; 2.
 DR ProDom; PD186100; IF2; 1.
 DR TIGRFAMs; TIGR00487; IF-2; 1.
 DR TIGRFAMs; TIGR00231; small_GTP; 1.

DR Pfam; PF04760; IF2_N; 2.
 DR ProDom; PD186100; IF2; 1.
 DR TIGRFAMs; TIGR00487; IF-2; 1.
 DR TIGRFAMs; TIGR00231; small_GTP; 1.
 DR PROSITE; PS01176; IF2; 1.
 KW Initiation factor; Protein biosynthesis; GTP-binding;
 KW Complete proteome.
 FT DOMAIN 343 491 G-DOMAIN.
 FT NP_BIND 349 356 GTP (BY SIMILARITY).
 FT NP_BIND 395 399 GTP (BY SIMILARITY).
 FT NP_BIND 449 452 GTP (BY SIMILARITY).
 SQ SEQUENCE 840 AA; 90912 MW; 60A69F798C140055 CRC64;

Query Match 93.1%; Score 27; DB 1; Length 840;
 Best Local Similarity 83.3%; Pred. No. 96;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GHKXKL 6
 |||||
 Db 313 GHKVKL 318

RESULT 7
 IF2_WIGBR
 ID IF2_WIGBR STANDARD; PRT; 841 AA.
 AC Q8D2X6;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Translation initiation factor IF-2.
 GN INFB OR WIGBR2260.
 OS Wigglesworthia glossinidia brevipalpis.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Wigglesworthia.
 OX NCBI_TaxID=36870;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22297718; PubMed=12219091;
 RA Akman L., Yamashita A., Watanabe H., Oshima K., Shiba T., Hattori M.,
 RA Alexoy S.;
 RT "Genome sequence of the endocellular obligate symbiont of tsetse
 RT flies, Wigglesworthia glossinidia.";
 RL Nat. Genet. 32:402-407(2002).
 CC -!- FUNCTION: One of the essential components for the initiation of
 CC protein synthesis. Protects formylmethionyl-tRNA from spontaneous
 CC hydrolysis and promotes its binding to the 30S ribosomal subunits.
 CC Also involved in the hydrolysis of GTP during the formation of the
 CC 70S ribosomal complex (By similarity).
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -!- SIMILARITY: Belongs to the IF-2 family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR ENBL; AE063521; BAC24372.1; -.
 DR HAMAP; MF_00100; -. 1.
 DR InterPro; IPR000795; EF_GTPbind.
 DR InterPro; IPR004161; EFTU_D2.
 DR InterPro; IPR000178; IF2.
 DR InterPro; IPR006847; IF2 N.
 DR InterPro; IPR005225; Small GTP.
 DR InterPro; IPR009000; Translat_factor.
 DR Pfam; PF03144; GTP_EFTU; 1.
 DR Pfam; PF04760; IF2_N; 2.
 DR ProDom; PD186100; IF2; 1.
 DR TIGRFAMs; TIGR00487; IF-2; 1.
 DR TIGRFAMs; TIGR00231; small_GTP; 1.

```

DR PROSITE; PS01176; IP2; FALSE NEG.
KW Initiation factor; Protein biosynthesis; GTP-binding;
KW Complete proteome.
FT DOMAIN 344 492 G-DOMAIN.
FT NP_BIND 350 357 GTP (BY SIMILARITY).
FT NP_BIND 396 400 GTP (BY SIMILARITY).
FT NP_BIND 450 453 GTP (BY SIMILARITY).
SQ SEQUENCE 841 AA; 94758 MW; B693CC2705D050F7 CRC64;

Query Match
Best Local Similarity 93.1%; Score 27; DB 1; Length 841;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GHKXKL 6
DB 313 GHKVKL 318

RESULT 8
SVL_METVA
ID SVL_METVA STANDARD; PRT; 878 AA.
AC Q58413;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Valyl-tRNA synthetase (EC 6.1.1.9) (Valine--tRNA ligase) (ValRS).
GN VALS OR MLI007.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcaceae; Methanocaldococcus.
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337993; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., FitzGerald L.M., Clayton R.A., Gocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA Uterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RA "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii."
RL Science 273:1058-1073 (1996).
CC -1- CATALYTIC ACTIVITY: ATP + L-valine + tRNA(Val) = AMP + diphosphate
CC + L-valyl-tRNA(Val).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: Belongs to class-I aminoacyl-tRNA synthetase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U67543; AAB99009.1; -.
DR PIR; F64425; F64425.
DR HSP; P96142; IGAX.
DR TIGR; M1007; -.
DR InterPro; IPR002300; tRNA-synt_1a.
DR InterPro; IPR001412; tRNA-synt_1.
DR InterPro; IPR002303; tRNA-synt_val.
DR InterPro; IPR009008; ValRS_1leRS_edit.
DR Pfam; PF00133; tRNA-synt_1; 1.
DR PRINTS; PR00986; TRNASYNTHAL.
DR TIGRFAMs; TIGR00422; vals; 1.
DR PROSITE; PS01176; AA_TRNA_LIGASE_I; 1.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KW Complete proteome.

```

```

FT SITE 43 53 "HIGH" REGION.
FT SITE 527 531 "KMSKS" REGION.
FT BINDING 530 530 ATP (BY SIMILARITY).
SQ SEQUENCE 878 AA; 102571 MW; 4D78BCA0DF2022E CRC64;

Query Match
Best Local Similarity 83.3%; Score 27; DB 1; Length 878;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GHKXKL 6
DB 247 GHKVKL 252

RESULT 9
MSI2_ARATH
ID MSI2_ARATH STANDARD; PRT; 415 AA.
AC O22458;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE WD-40 repeat protein MSI2.
GN MSI2 OR AT2G16780 OR T24121.19.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97480095; PubMed=9338962;
RA Ach R.A., Taranto P., Gruissem W.;
RA "A conserved family of WD-40 proteins binds to the retinoblastoma
RT protein in both plants and animals."
RL Plant Cell 9:1595-1606 (1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=20083487; PubMed=10617197;
RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblyum T.V.,
RA Buell C.R., Ketchum K.A., Lee J.J., Renning C.M., Koo H.L.,
RA Moffat K.S., Cronin L.A., Shen M., Pal G., Van Aken S., Umayam L.,
RA Tallon L.J., Gill J.E., Adams M.D., Carrera A.J., Creasy T.H.,
RA Goodman H.M., Somerville C.R., Coppenhaver G.P., Preuss D.,
RA Nierman W.C., White O., Eisen J.A., Salzberg S.L., Fraser C.M.,
RA Venter J.C.;
RA "Sequence and analysis of chromosome 2 of the plant Arabidopsis
RT thaliana."
RL Nature 402:761-768 (1999).
CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -1- SIMILARITY: Belongs to the WD-repeat RBAP46/RBAP48/MSI1 family.
CC -1- SIMILARITY: Contains 5 WD repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF016847; AAE70243.1; -.
DR EMBL; AC005825; AAD24611.1; -.
DR PIR; B84544; B84544.
DR InterPro; IPR001680; WD40.
DR Pfam; PF00400; WD40; 5.
DR PRINTS; PR00320; GPROTEINBRPT.
DR SMART; SM00320; WD40; 5.
DR PROSITE; PS00678; WD_REPEATS_1; 2.
DR PROSITE; PS50082; WD_REPEATS_2; 5.
DR PROSITE; PS50294; WD_REPEATS_REGION; 1.
KW Repeat; WD repeat; Nuclear protein.

```

```

FT REPEAT 166 206 WD 1.
FT REPEAT 215 255 WD 2.
FT REPEAT 258 298 WD 3.
FT REPEAT 302 342 WD 4.
FT REPEAT 361 401 WD 5.
SQ SEQUENCE 415 AA; 46704 MW; D89F99627C5D84E CRC64;

Query Match 89.7%; Score 26; DB 1; Length 415;
Best Local Similarity 66.7%; Pred. No. 78;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GHKXKL 6
DB 361 GHKAKI 366

RESULT 10
ID MS13 ARATH STANDARD; PRT; 424 AA.
AC Q22469; O49612;
DT 15-JUL-1998 (Rel. 36, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE WD-40 repeat protein MS13.
GN MS13 OR A4G35050 OR M4E13.110.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]_
RP SEQUENCE FROM N.A.
RX MEDLINE=97480095; PubMed=9338962;
RA Ach R.A., Taranto P., Gruissem W.;
RT "A conserved family of WD-40 proteins binds to the retinoblastoma
RT protein in both plants and animals.";
RL Plant Cell 9:1595-1606 (1997).
RN [2]_
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=20083488; PubMed=10617198;
RA Mayer K.F.X., Schueller C., Wambutt R., Murphy G., Volckaert G.,
RA Pohl T., Dueterhoeft A., Stiekema W., Entian K.-D., Terry N.,
RA Harris B., Ansoorge W., Brandt P., Grivell L.A., Rieger M.,
RA Weichselgartner M., de Simone V., Obermaier B., Mache R., Mueller M.,
RA Kreis M., Delisny M., Puigdomenech P., Watson M., Schmidtheini T.,
RA Reichert B., Portetelle D., Perez-Alonso M., Boutry M., Bancroft I.,
RA Vos P., Hehseisel J., Zimmermann W., Wedler H., Ridley P.,
RA Langham S.-A., McCullagh B., Bilham L., Robben J.,
RA Van der Schueren J., Grymonprez B., Chuang Y.-J., Vandenbussche F.,
RA Braeken M., Weltjens I., Voet M., Bastiaens I., Aert R., Defoor E.,
RA Weitzenecker T., Bothe G., Ramsperger U., Hilbert H., Braun M.,
RA Holzer E., Brandt A., Peters S., van Staveren M., Dirkse W.,
RA Moeljan P., Klein Lankhorst R., Rose M., Hauf J., Koetter P.,
RA Bernreiser S., Hempel S., Rehdauisch M., Lamberth S., Van den Daele H.,
RA De Keyser A., Buyschaert C., Gielen J., Villarroel R., De Clercq R.,
RA Van Montagu M., Rogers J., Cronin A., Quail M.A., Bray-Allen S.,
RA Clark L., Doggett J., Hall S., Kay M., Leonard N., McIay K., Mayes R.,
RA Borkova D., Bloeker H., Scharfe M., Grimm M., Loehner T.-H.,
RA Dose S., de Haan M., Maarse A.C., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Farman B., Granderath K., Dauner D., Herzi A.,
RA Neumann S., Argiriou A., Vitale D., Liguori R., Piravandi E.,
RA Massenet O., Quigley F., Clabaud G., Muendlein A., Felber R.,
RA Schnabl S., Hiller R., Schmidt W., Lecharny A., Aubourg S.,
RA Chedford F., Cooke R., Berger C., Monfort A., Casacuberta E.,
RA Gibbons T., Weber N., Vandenbol M., Bagues M., Terol J., Torres A.,
RA Perez-Perez A., Purnelle B., Bent E., Johnson S., Tacon D., Jesse T.,
RA Heijnen L., Schwarz S., Scholler P., Heber S., Francis P., Bielke C.,
RA Frishman D., Haase D., Lemcke K., Mewes H.-W., Stocker S.,
RA Zaccaria P., Wilson R.K., de la Bastide M., Habermann K.,
RA Parnell L., Dedhia N., Gnoj L., Schutz K., Huang E., Spiegel L.,
RA Sekhon M., Murray J., Sheet P., Cordes M., Abu-Threideh J.,

```

```

RA Stoneking T., Kalicki J., Graves T., Harmon G., Edwards J.,
RA Latreille P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,
RA Mix P., Bentley D., Fulton B., Miller N., Greco T., Kemp K.,
RA Kramon J., Fulton L., Mardis E., Dantes M., Pepin K., Hillier L.W.,
RA Nelson J., Spieth J., Ryan E., Andrews S., Geisel C., Layman D.,
RA Du H., Ali J., Berghoff A., Jones K., Drone K., Cotton M., Joshi C.,
RA Antoniou B., Zidanic M., Strong C., Sun H., Lamar B., Yordan C.,
RA Ma P., Zhong J., Preston R., Vil D., Shekier M., Matero A., Shah R.,
RA Swaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Tili S.,
RA Granat S., Shohdy N., Hasegawa A., Hamed A., Lodhi M., Johnson A.,
RA Chen E., Marra M.A., Martienssen R., McCombie W.R.;
RT "Sequence and analysis of chromosome 4 of the plant Arabidopsis
RT thaliana.";
RL Nature 402:769-777 (1999).
CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -!- SIMILARITY: Belongs to the WD-repeat RBP46/RBP48/MS11 family.
CC -!- SIMILARITY: Contains 5 WD repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF016848; AAB70344.1; -;
DR EMBL; AL022023; CNA17770.1; -;
DR EMBL; AL161586; CAB90222.1; -;
DR PIR; T05775; T05775.
DR InterPro; IPR001680; WD40.
DR Pfam; PF00400; WD40; 5.
DR PRINTS; PR00320; GPROTEINRPT.
DR SMART; SM00320; WD40; 5.
DR PROSITE; PS00678; WD_REPEATS_1; 2.
DR PROSITE; PS00882; WD_REPEATS_2; 4.
DR PROSITE; PS0294; WD_REPEATS_REGION; 1.
KW Repeat; WD repeat; Nuclear protein.
FT REPEAT 167 207 WD 1.
FT REPEAT 216 256 WD 2.
FT REPEAT 259 299 WD 3.
FT REPEAT 303 343 WD 4.
FT REPEAT 362 402 WD 5.
FT CONFLICT 223 224 DV -> EL (IN RBF. 1).
SQ SEQUENCE 424 AA; 47983 MW; D83E2B4913468A0A CRC64;

Query Match 89.7%; Score 26; DB 1; Length 424;
Best Local Similarity 66.7%; Pred. No. 80;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GHKXKL 6
DB 362 GHKAKI 367

RESULT 11
YAV5 XANCV STANDARD; PRT; 445 AA.
AC P19520;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Hypothetical 50 kDa avirulence protein in avrBs1 region.
OS Xanthomonas campestris (pv. vesicatoria).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=341;
RN [1]_
RP SEQUENCE FROM N.A.
RX MEDLINE=92145033; PubMed=2979910;
RA Ronald P.C., Staskawicz B.J.;
RT "The avirulence gene avrBs1 from Xanthomonas campestris pv.

```

```

RT vesicatoria encodes a 50-kD protein."
RL Mol. Plant Microbe Interact. 1:191-198(1998).
CC -!- MISCELLANEOUS: THIS IS ONE OF THE HYPOTHETICAL PROTEINS CODED BY
CC THE OPEN READING FRAMES WITHIN THE REGION REQUIRED FOR AVIRUS
CC ACTIVITY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M32142; AAA27594.1; -.
CC KW Hypothetical protein; Plasmid; Virulence.
CC SEQUENCE 445 AA; 49799 MW; 381BA3E290EA175F CRC64;
SQ
Query Match 89.7%; Score 26; DB 1; Length 445;
Best Local Similarity 66.7%; Pred. No. 84;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 GHXXKL 6
Db 342 GHKSKM 347
|||||
-----
RESULT 12
IF2_COXBU
ID IF2_COXBU STANDARD; PRT; 803 AA.
AC Q83BS1;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DE Translation initiation factor IF-2.
GN INF6 OR CBU1432
OS Coxiella burnetii.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Legionellales;
OC Coxiellaceae; Coxiella.
OX NCBI_TaxID=777;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Nine Mile Phase I / RSA 493;
RX MEDLINE=2260857; PubMed=12704232;
RA Seshadri R., Paulsen I.T., Eisen J.A., Read T.D., Nelson K.E.,
RA Nelson W.C., Ward N.L., Tettelin H., Daviden T.M., Beanan M.J.,
RA DeBoy R.T., Daugherty S.C., Brinkac L.M., Madupu R., Dodson R.J.,
RA Khouri H.M., Lee K.H., Carty H.A., Scanlan D., Heinzen R.A.,
RA Thompson H.A., Samuel J.B., Fraser C.M., Heidelberg J.F.;
RT "Complete genome sequence of the Q-fever pathogen, Coxiella
RT burnetii."
RL Proc. Natl. Acad. Sci. U.S.A. 100:5455-5460(2003).
CC -!- FUNCTION: One of the essential components for the initiation of
CC protein synthesis. Protects formylmethionyl-tRNA from spontaneous
CC hydrolysis and promotes its binding to the 30S ribosomal subunits.
CC also involved in the hydrolysis of GTP during the formation of the
CC 70S ribosomal complex (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to the IF-2 family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; A5016964; AA090929.1; -.
CC TIGR; CEU1432; -.
CC HAMAP; MF 00100; -.
CC InterPro; IPR004161; EFTU D2.
CC InterPro; IPR000795; EF_GTPbind.

```

```

DR InterPro; IPR000178; IF2.
DR InterPro; IPR006847; IF2_N.
DR InterPro; IPR000717; PCI.
DR InterPro; IPR005225; Small_GTP.
DR Pfam; PF00009; GTP_EFTU; 1.
DR Pfam; PF03144; GTP_EFTU_D2; 2.
DR Pfam; PF04760; IF2_N_2.
DR ProDom; PD186100; IF2; 1.
DR TIGRFAMs; TIGR00487; IF-2; 1.
DR TIGRFAMs; TIGR00231; small_GTP; 1.
DR PROSITE; PS01176; IF2; 1.
KW Initiation factor; Protein biosynthesis; GTP-binding;
KW Complete proteome.
KW DOMAIN 305 453 G-DOMAIN.
FT NP_BIND 311 318 GTP (BY SIMILARITY).
FT NP_BIND 357 361 GTP (BY SIMILARITY).
FT NP_BIND 411 414 GTP (BY SIMILARITY).
SQ SEQUENCE 803 AA; 88485 MW; 00C73115B6B3D5C9 CRC64;
Query Match 89.7%; Score 26; DB 1; Length 803;
Best Local Similarity 83.3%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GHXXKL 6
Db 274 GHKPKL 279
|||||
-----
RESULT 13
RPB1_PLAFD
ID RPB1_PLAFD STANDARD; PRT; 2452 AA.
AC P14248;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE DNA-directed RNA polymerase II largest subunit (EC 2.7.7.6).
GN RPB1.
OS Plasmodium falciparum (isolate CDC / Honduras).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5836;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9098932; PubMed=2690004;
RA Li W.B., Baik D.J., Gu H., Tanaka M., Fox B.A., Inselburg J.;
RT "An enlarged largest subunit of Plasmodium falciparum RNA polymerase
RT II defines conserved and variable RNA polymerase domains."
RL Nucleic Acids Res. 17:9621-9636(1989).
CC -!- FUNCTION: DNA-dependent RNA polymerase catalyzes the transcription
CC of DNA into RNA using the four ribonucleoside triphosphates as
CC substrates.
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC (RNA)(N).
CC -!- SUBUNIT: EACH CLASS OF RNA POLYMERASE IS ASSEMBLED FROM TEN TO
CC TWELVE DIFFERENT POLYPEPTIDES. THIS POLYPEPTIDE IS THE LARGEST
CC COMPONENT OF RNA POLYMERASE II.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- PTM: The tandem 7 residues repeats can be highly phosphorylated.
CC -!- MISCELLANEOUS: Three distinct zinc-containing RNA polymerases were
CC found in eukaryotic nuclei: polymerase I for the ribosomal RNA
CC precursor, polymerase II for the mRNA precursor, and polymerase
CC III for 5S and tRNA genes.
CC -!- SIMILARITY: Belongs to the RNA polymerase beta' chain family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----

```

```

DR EMBL; X16561; CAA34560.1; -.
DR PIR; S07485; RNQ22L.
DR InterPro; IPR00722; RNA_pol_A.
DR InterPro; IPR007080; RNA_pol_Rpb1_1.
DR InterPro; IPR007066; RNA_pol_Rpb1_3.
DR InterPro; IPR007083; RNA_pol_Rpb1_4.
DR InterPro; IPR007081; RNA_pol_Rpb1_5.
DR InterPro; IPR007075; RNA_pol_Rpb1_6.
DR InterPro; IPR007073; RNA_pol_Rpb1_7.
DR InterPro; IPR006592; RNA_pol_N.
DR InterPro; IPR00684; RNA_pol_II_repeat.
DR Pfam; PF04997; RNA_pol_Rpb1_1; 1.
DR Pfam; PF06623; RNA_pol_Rpb1_2; 1.
DR Pfam; PF04983; RNA_pol_Rpb1_3; 1.
DR Pfam; PF05000; RNA_pol_Rpb1_4; 1.
DR Pfam; PF04998; RNA_pol_Rpb1_5; 1.
DR Pfam; PF04992; RNA_pol_Rpb1_6; 1.
DR Pfam; PF04990; RNA_pol_Rpb1_7; 1.
DR Pfam; PF05001; RNA_pol_Rpb1_8; 12.
DR SMART; SM00663; RPOA_N; 1.
DR PROSITE; PS00115; RNA_POL_II_REPEAT; 9.
KW Transferase; DNA-directed RNA polymerase; Transcription; Zinc; Repeat;
KW DNA-binding; Nuclear protein; Phosphorylation; Zinc-finger.
FT ZN_FING 68 84
FT DNA_BIND 378 411
FT DOMAIN 707 725
FT DOMAIN 1093 1128
FT DOMAIN 1144 1159
FT DNA_BIND 1182 1193
FT DOMAIN 1687 1694
FT DOMAIN 1258 1290
FT DOMAIN 1261 1290
FT DOMAIN 1602 1612
FT DOMAIN 1746 1759
FT DOMAIN 1806 1820
FT DOMAIN 2061 2246
FT DOMAIN 2247 2384
SQ SEQUENCE 2452 AA; 278166 MW; F995E117F617A48F CRC64;

Query Match 89.7%; Score 26; DB 1; Length 2452;
Best Local Similarity 66.7%; Pred. No. 5.3e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GHKKKL 6
DB 490 GHKAKI 495

RESULT 14
RL22_HALMA
ID RL22_HALMA STANDARD; PRT; 154 AA.
AC P10970;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE S08 ribosomal protein L22P (Hmal22) (HL23).
GN RPL22P.
OS Haloarcula marismortui (Halo bacterium marismortui).
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC Halobacteriaceae; Haloarcula.
CX NCBI_TaxID=2238;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90153945; PubMed=2406244;
RA Arndt E., Kremer W., Hatakeyama T.;
RT "Organization and nucleotide sequence of a gene cluster coding for
RT eight ribosomal proteins in the archaeobacterium Halo bacterium
RT marismortui.";
RL J. Biol. Chem. 265:3034-3039 (1990).
RN [2]
RP SEQUENCE.
RX MEDLINE=69052886; PubMed=3191994;
RA Hatakeyama T., Hatakeyama T., Kimura M.;

```

```

RT "The primary structures of ribosomal proteins L16, L23 and L33 from
RT the archaeobacterium Halobacterium marismortui.";
RL FEBS Lett. 240:21-28 (1988).
RN [3]
RP SEQUENCE OF 1-26.
RX MEDLINE=89062418; PubMed=3196689;
RA Walsh M.J., McDougall J., Wittmann-Liebold B.;
RT "Extended N-terminal sequencing of proteins of archaeobacterial
RT ribosomes blotted from two-dimensional gels onto glass fiber and
RT poly(vinylidene difluoride) membrane.";
RL Biochemistry 27:6867-6876 (1988).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS).
RC STRAIN=ATCC 43049;
RX MEDLINE=20396344; PubMed=10937989;
RA Pan N., Nissen P., Hansen J., Moore P.B., Steitz T.A.;
RT "The complete atomic structure of the large ribosomal subunit at 2.4
RT A resolution.";
RL Science 289:905-920 (2000).
CC -!- SIMILARITY: Belongs to the L22P family of ribosomal proteins.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J05222; AAA86864.1; -.
DR PIR; H35063; R5HS22.
DR PDB; 1FFK; 26-SEP-01.
DR PDB; 1K8A; 19-JUL-02.
DR PDB; 1K9M; 19-JUL-02.
DR PDB; 1K0L; 19-JUL-02.
DR PDB; 1MLK; 23-AUG-02.
DR PDB; 1M90; 06-SEP-02.
DR InterPro; IPR001063; Ribosomal_L22.
DR InterPro; IPR005721; Ribosomal_L22_17.
DR Pfam; PF00237; Ribosomal_L22; 1.
DR ProDom; PD001032; Ribosomal_L22; 1.
DR TrGFams; TIGR01038; L22_arch; 1.
DR PROSITE; PS00464; RIBOSOMAL_L22; 1.
KW Ribosomal protein; 3D-structure.
FT INIT_MET 0
SQ SEQUENCE 154 AA; 16810 MW; 86EA22CFF448907C CRC64;

Query Match 86.2%; Score 25; DB 1; Length 154;
Best Local Similarity 66.7%; Pred. No. 47;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GHKKKL 6
DB 67 GHKSKV 72

RESULT 15
VIRF_SHIFL
ID VIRF_SHIFL STANDARD; PRT; 262 AA.
AC Q04248; Q9AFW5;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Virulence regulon transcriptional activator virF.
GN VIRF OR CP0046.
OS Shigella flexneri,
OS Shigella dysenteriae, and
OS Shigella sonnei.
OG Plasmid pWR100, Plasmid pMYSH6000, and Plasmid pCP301.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
CX NCBI_TaxID=623, 622, 624;
RN [1]

```


RP SEQUENCE FROM N.A.
 RC SPECIES=S.flexneri; STRAIN=M90T / Serotype 5a; PLASMID=pWR100;
 RX MEDLINE=20566792; PubMed=11151111;
 RA Buchrieser C., Glaser P., Rusnock C., Nedjari H., d'Hauteville H.,
 RA Kunst F., Sansonetti P., Parsot C.;
 RT "The virulence plasmid pWR100 and the repertoire of proteins secreted
 RT by the type III secretion apparatus of Shigella flexneri.";
 RL Mol. Microbiol. 38:760-771(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC SPECIES=S.flexneri; STRAIN=M90T / Serotype 5; PLASMID=pWR100;
 RX MEDLINE=21189246; PubMed=11202750;
 RA Venkatesan M.M., Goldberg M.B., Rose D.J., Grotbeck E.J., Burland V.,
 RA Blattner F.R.;
 RT "Complete DNA sequence and analysis of the large virulence plasmid of
 RT Shigella flexneri.";
 RL Infect. Immun. 69:3271-3285(2001).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC SPECIES=S.flexneri; STRAIN=YSH6000 / Serotype 2a;
 RX MEDLINE=87032409; PubMed=3021627;
 RA Sakai T., Sasaki K., Makino S., Yoshikawa M.;
 RT "DNA sequence and product analysis of the virF locus responsible for
 RT Congo red binding and cell invasion in Shigella flexneri 2a.";
 RL Infect. Immun. 54:395-402(1986).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC SPECIES=S.flexneri; STRAIN=301 / Serotype 2a; PLASMID=pCF301;
 RX MEDLINE=2272406; PubMed=12384590;
 RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
 RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
 RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
 RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
 RA Yu J.;
 RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
 RT through comparison with genomes of Escherichia coli K12 and O157.";
 RL Nucleic Acids Res. 30:4432-4441(2002).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC SPECIES=S.dysenteriae; STRAIN=CG097;
 RA Yao R., Reddy L.V., Palchaudhuri S.;
 RL Submitted (MAR-1991) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SEQUENCE FROM N.A.
 RC SPECIES=S.sonnei;
 RX MEDLINE=89212881; PubMed=2651305;
 RA Kato J.I., Ito K.I., Nakamura A., Watanabe H.;
 RT "Cloning of regions required for contact hemolysis and entry into
 RT LLC-MK2 cells from Shigella sonnei form I plasmid: virF is a
 RT positive regulator gene for these phenotypes.";
 RL Infect. Immun. 57:1391-1398(1989).
 RN [7]
 RP SIMILARITY TO ARAC FAMILY.
 RX MEDLINE=92326642; PubMed=1625585;
 RA Dorman C.J.;
 RT "The VirF protein from Shigella flexneri is a member of the AraC
 RT transcription factor superfamily and is highly homologous to Rns, a
 RT positive regulator of virulence genes in enterotoxigenic Escherichia
 RT coli.";
 RL Mol. Microbiol. 6:1575-1575(1992).
 CC -!- FUNCTION: Transcriptional activator of the virB gene which is
 CC itself an activator of the ipaABCD virulence regulon.
 CC -!- SIMILARITY: BELONGS TO THE ARAC/XYLIS FAMILY OF TRANSCRIPTIONAL
 CC REGULATORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; AL391753; CAC05780.1; -
 DR EMBL; AF348706; AAK18362.1; ALT_INIT.
 DR EMBL; M29172; AAA26545.1; -
 DR EMBL; X16661; CAA34648.1; -
 DR EMBL; AF386526; AAL72294.1; ALT_INIT.
 DR EMBL; X58464; CAA41378.1; -
 DR PIR; A47605; A47605.
 DR PIR; A60105; A60105.
 DR PIR; S14646; S14646.
 DR InterPro; IPR000005; HTHArac.
 DR Pfam; PF00165; HTH_Arac; 2.
 DR PRINTS; PR00032; HTHARAC.
 DR SMART; SMO0342; HTH_ARAC; 1.
 DR PROSITE; PS00041; HTH_ARAC_FAMILY_1; 1.
 DR PROSITE; PS01124; HTH_ARAC_FAMILY_2; 1.
 KW Plasmid; Virulence; Transcription regulation; Activator; DNA-binding;
 KW Complete proteome.
 FT DNA_BIND 177 196 H-T-H MOTIF (BY SIMILARITY).
 SQ SEQUENCE 262 AA; 30549 MW; 3CB18BA52DE53443 CRC64;
 Query Match 86.2%; Score 25; DB 1; Length 262;
 Best Local Similarity 66.7%; Pred.No. 83;
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 GHKKKL 6
 Db 5 GHKKKI 10
 Search completed: March 4, 2004, 13:08:09
 Job time : 4.38235 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2004, 13:02:03 ; Search time 11.2941 Seconds
(without alignments)
167.619 Million cell updates/sec

Title: SEQ-D

Perfect score: 29

Sequence: 1 ghxkl 6

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

- 1: sp archaea.*
- 2: sp bacteria.*
- 3: sp fungi.*
- 4: sp human.*
- 5: sp invertebrate.*
- 6: sp mammal.*
- 7: sp mhc.*
- 8: sp organelle.*
- 9: sp phage.*
- 10: sp plant.*
- 11: sp rodent.*
- 12: sp virus.*
- 13: sp vertebrate.*
- 14: sp unclassified.*
- 15: sp virus.*
- 16: sp bacterioph.*
- 17: sp archaea.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	28	96.6	139	12 Q9IKD0	Q9ikd0 rat sialoda
2	28	96.6	153	16 Q9KUG8	Q9kug8 vibrio chol
3	28	96.6	188	16 Q7UR39	Q7ur39 rhodopirell
4	28	96.6	211	10 Q94GR2	Q94gr2 oryza sativ
5	28	96.6	218	10 P93007	P93007 arabidopsis
6	28	96.6	259	5 Q9XTD3	Q9xtd3 caenorhabdi
7	28	96.6	304	10 Q80942	Q80942 arabidopsis
8	28	96.6	397	10 Q941R1	Q941r1 lycopersico
9	28	96.6	407	4 Q8N7X4	Q8n7x4 homo sapien
10	28	96.6	522	15 Q9PXU3	Q9pxu3 chimpanzee
11	28	96.6	523	15 P89903	P89903 chimpanzee
12	28	96.6	786	5 P91451	P91451 caenorhabdi
13	28	96.6	867	13 Q9ET70	Q9et70 xenopus lae
14	27	93.1	103	16 Q9KML9	Q9kml9 vibrio chol
15	27	93.1	123	16 Q82U99	Q82u99 nitrosomona
16	27	93.1	170	16 Q06613	Q06613 mycobacteri

17	27	93.1	170	16 Q7VEX9	Q7vex9 mycobacteri
18	27	93.1	172	11 Q9D498	Q9d4s8 mus musculu
19	27	93.1	177	5 Q61695	Q61695 manduca sex
20	27	93.1	177	16 P71948	P71948 mycobacteri
21	27	93.1	202	11 Q7TS87	Q7ts87 mus musculu
22	27	93.1	236	8 Q9G4C8	Q9g4c8 thraustochy
23	27	93.1	288	16 Q67299	Q67299 aquifex aeo
24	27	93.1	297	5 Q812F4	Q812f4 plasmodium
25	27	93.1	300	17 Q87T25	Q87t25 methanosarc
26	27	93.1	302	16 Q8K611	Q8k611 streptococc
27	27	93.1	332	5 Q86F84	Q86f84 schistosoma
28	27	93.1	341	9 Q8W750	Q8w750 bacterioph
29	27	93.1	383	11 Q8VDJ1	Q8vdj1 mus musculu
30	27	93.1	384	16 Q9HTH4	Q9hth4 pseudomonas
31	27	93.1	391	16 Q88GZ4	Q88gz4 pseudomonas
32	27	93.1	392	4 Q9H5N3	Q9h5n3 homo sapien
33	27	93.1	443	4 Q96F64	Q96f64 homo sapien
34	27	93.1	473	2 Q51898	Q51898 borrelia af
35	27	93.1	482	16 Q51783	Q51783 borrelia bu
36	27	93.1	513	5 Q960L1	Q960l1 drosophila
37	27	93.1	551	10 Q9LJF1	Q9ljf1 arabadopsis
38	27	93.1	605	16 Q7TU36	Q7tu36 prochloroco
39	27	93.1	726	4 Q8NBF2	Q8nbf2 homo sapien
40	27	93.1	737	4 Q86U99	Q86u99 homo sapien
41	27	93.1	777	12 Q83073	Q83073 leishmania
42	27	93.1	816	5 Q8IU20	Q8iu20 branchiosto
43	27	93.1	842	4 Q9Y6P5	Q9y6f5 homo sapien
44	27	93.1	874	4 Q7Z644	Q7z644 homo sapien
45	27	93.1	960	11 Q7TSG2	Q7tsg2 mus musculu

ALIGNMENTS

RESULT 1

ID	Q9IKD0	PRELIMINARY;	PRT;	139 AA.
AC	Q9IKD0;			
DT	01-OCT-2000 (Tremblrel. 15, Created)			
DT	01-OCT-2000 (Tremblrel. 15, Last sequence update)			
DT	01-JUN-2003 (Tremblrel. 24, Last annotation update)			
DE	15 kDa nonstructural protein.			
OS	Rat sialodacryoadenitis coronavirus.			
OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;			
OC	Coronaviridae; Coronavirus.			
OX	NCBI_TaxID=92931;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=SDAV-681;			
RX	MEDLINE=20342543; PubMed=1082653;			
RA	Yoo D., Pei Y., Christie N., Cooper M.;			
RT	"Primary Structure of the Sialodacryoadenitis Virus Genome: Sequence of the Structural-Protein Region and its Application for Differential			
RT	Diagnosis";			
RL	Clin. Diagn. Lab. Immunol. 7:568-573 (2000).			
DR	EMBL; AF207551; AAF97739.1; -			
DR	GO; GO:0019012; C:virion; IEA.			
DR	InterPro; IPR005603; Corona_NS4.			
KW	Pfam; PF03905; Corona_NS4; 1.			
KW	Nonstructural protein.			
SQ	SEQUENCE 139 AA; 15387 MW; A4E466B6DE31FFEB CRC64;			

Query Match 96.6%; Score 28; DB 12; Length 139;
Best Local Similarity 83.3%; Pred. No. 73;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GHXKL 6
|||
Db 5 GEXKL 10

RESULT 2

Q9KUG8

ID Q9KUG8 PRELIMINARY; PRT; 153 AA.
 AC Q9KUG8;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE Hypothetical protein VC0553.
 GN VC0553.
 OS Vibrio cholerae.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
 OC Vibrionaceae; Vibrio.
 OX NCBI_TaxID=666;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=El Tor N16961 / Serotype O1;
 RX MEDLINE=20406833; PubMed=10952301;
 RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
 RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
 RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
 RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
 RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
 RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
 RA Fraser C.M.;
 RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
 cholerae.";
 RL Nature 406:477-483(2000).
 DR EMBL; AF004141; AAF93721.1; -.
 DR PIR; B82309; B82309.
 DR TIGR; VC0553; -.
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0005524; P:ATP binding; IEA.
 DR GO; GO:0004009; P:ATP-binding cassette (ABC) transporter acti. . .; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR003439; ABC transporter.
 DR PROSITE; PS00211; ABC TRANSPORTER 1; 1.
 DR Hypothetical protein: Complete proteome.
 KW Hypothetical protein: Complete proteome.
 SQ SEQUENCE 153 AA; 16761 MW; 1D9E52C300A6251B CRC64;

Query Match 96.6%; Score 28; DB 16; Length 153;
 Best Local Similarity 83.3%; Pred. No. 81;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GHXXKL 6
 DB 89 GHKAKL 94

RESULT 3
 Q7UR39
 ID Q7UR39 PRELIMINARY; PRT; 188 AA.
 AC Q7UR39;
 DT 01-OCT-2003 (TREMBlrel. 25, Created)
 DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE Hypothetical protein.
 GN R55907.
 OS Rhodospirillum rubrum.
 OC Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales;
 OC Planctomycetaceae; Pirellula.
 OX NCBI_TaxID=117;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=1;
 RX MEDLINE=22735913; PubMed=12835416;
 RA Gloeckner F.O., Kube M., Bauer M., Teeling H., Lombardot T.,
 RA Ludwig W., Gade D., Beck A., Borzym K., Heitmann K., Rabus R.,
 RA Schlesner H., Anann R., Reinhardt R.;
 RT "Complete genome sequence of the marine planctomycete Pirellula sp.
 strain 1.";
 RT Proc. Natl. Acad. Sci. U.S.A. 100:8298-8303(2003).
 RL EMBL; EX294143; CAD74503.1; -.
 KW Hypothetical protein: Complete proteome.
 SQ SEQUENCE 188 AA; 19745 MW; 82E552867FC8B3BD CRC64;

Query Match 96.6%; Score 28; DB 16; Length 188;
 Best Local Similarity 83.3%; Pred. No. 99;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 GHXXKL 6
 DB 34 GHKSKL 39

RESULT 4
 Q94GR2
 ID Q94GR2 PRELIMINARY; PRT; 211 AA.
 AC Q94GR2;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE Hypothetical protein.
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzoideae; Oryza.
 OX NCBI_TaxID=4530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Nipponbare;
 RA Buell C.R., Yuan Q., Ouyang S., Moffat K.S., Hill J.N., Gansberger K.,
 RA Brenner M., Burgess S., Hance M., Shvartsbeyn M., Tsitrin T.,
 RA Riggs F., Heiao J., Zismann V., Blunt S., Pai G., VanAken S.E.,
 RA Utterback T.R., Feldblyum T.V., Quackenbush J., Salzberg S.L.,
 RA White O., Fraser C.M.;
 RT "Oryza sativa chromosome 3 BAC OSJNB0022E02 genomic sequence.";
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC087797; AAK50121.1; -.
 DR Gramene; Q94GR2; -.
 DR Hypothetical protein.
 KW Hypothetical protein.
 SQ SEQUENCE 211 AA; 22539 MW; 5974D27F56E7AF2A CRC64;

Query Match 96.6%; Score 28; DB 10; Length 211;
 Best Local Similarity 83.3%; Pred. No. 11e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GHXXKL 6
 DB 36 GHKTKL 41

RESULT 5
 P93007
 ID P93007 PRELIMINARY; PRT; 218 AA.
 AC P93007;
 DT 01-MAY-1997 (TREMBlrel. 03, Created)
 DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
 DE Putative AP2 domain transcription factor.
 GN AT2G33710.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=20083487; PubMed=10617197;
 RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
 RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblyum T.V.,
 RA Buell C.R., Ketchum K.A., Lee J.J., Rensing C.M., Koo H., Moffat K.S.,
 RA Cronin L.A., Shen M., VanAken S.E., Umayam L., Tallon L.J., Gill J.E.,
 RA Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
 RA Copenhaver G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,
 RA Salzberg S.L., Fraser C.M., Venter J.C.;
 RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
 thaliana.";

RA Lightning J., Lloyd C., Mcmurray A., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifken L., Rocpra A., Saunders D., Showlken R.,
RA Samlson D., Smith A., Sonhammer E., Staden R., Sulston J.,
RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA Watson A., Weinstock L., Wilkinson-Sproat J., Woldman P.,
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of *C.
RT elegans*." ;
RL Nature 368:32-38 (1994).
RN [3]
RN SEQUENCE FROM N.A.
RP
RA White S.;
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z82077; CAB04941.1; -.
DR EMBL; Z82077; CAB04941.1; -.
DR EMBL; AL032650; CAB04941.1; JOINED.
DR EMBL; AL032650; CAA21705.1; -.
DR EMBL; Z82077; CAA21705.1; JOINED.
DR PIR; T26298; T26298.
DR Wormprep; W09C5.1; CE20164.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0005840; C:ribosome; IEA.
DR GO; GO:0003735; F:structural constituent of ribosome; IEA.
DR GO; GO:0006412; P:protein biosynthesis; IEA.
DR InterPro; IPR001047; Ribosomal S8e.
DR Pfam; PF01201; Ribosomal S8e; I.
DR Sfam; Sfam259 AA; 25717 MW; BFO50C6FB0AD2541 CRC64;
SQ
Query Match 96.6%; Score 28; DB 5; Length 259;
Best Local Similarity 83.3%; Pred.No.1.4e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps
QY 1 GHKXKL 6
||| ||
Db 44 GHKAKL 49
RESULT 7
080942
ID 080942 PRELIMINARY; PRT; 304 AA.
AC 080942;
DT DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Putative C2H2-type zinc finger protein.
GN AT2G37740.
OS Arabidopsis thaliana (Mouse-ear cross).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OC NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RC MEDLINE=20083487; PubMed=10617197;
EX Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldolym T.V.,
Buell C.R., Ketchum K.A., Lee J.J., Rensing C.M., Koo H., Moffat K.S.,
Cronin L.A., Shen M., Vanaken S.E., Umayam L., Tallon L.J., Gill J.B.,
Adams M.D., Carrera A.J., Crasay T.H., Goodman H.M., Somerville C.R.,
Copenhaver G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,
Salzberg S.L., Fraser C.M., Venter J.C.;
FT "Sequence and analysis of chromosome 2 of the plant Arabidopsi.
FT thaliana." ;
RL Nature 402:761-768 (1999).
RN [2]
RN SEQUENCE FROM N.A.
RP STRAIN=cv. Columbia;
RL Lin X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC004684; AAC23644.1; -.
DR FIR; T02540; T02540.
DR InterPro; IPR007087; Znf_C2H2.
DR Pfam; PF00036; zf-C2H2; I.
DR SMART; SM00355; Znf_C2H2; 1.

DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 1.
 DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 1.
 KW Metal-binding; Zinc; Zinc-finger.
 SQ SEQUENCE 304 AA; 34125 MW; 363FC671BBBA750 CRC64;

Query Match 96.6%; Score 28; DB 10; Length 304;
 Best Local Similarity 83.3%; Pred. No. 1.6e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GHXXKL 6
 |||||
 Db 206 GHKXKL 211

RESULT 8

Q941R1 ID Q941R1 PRELIMINARY; PRT; 397 AA.
 AC Q941R1
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DE 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
 DE Transaldolase TOTAL2
 OS Lycopersicon esculentum (Tomato).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
 OC Lamnids; Solanales; Solanaceae; Solanum.
 OX NCBI_TaxID=4081;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Callau M., Quick W.P.;
 RT "Tomato Transaldolase (TOTAL2).";
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY007225; AAG34725.1;
 DR GO; GO:0016228; F:aldolase activity; IEA.
 DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
 DR InterPro; IPR001585; P:carbohydrate metabolism; IEA.
 DR Pfam; PF00923; Transaldolase; 1.
 SQ SEQUENCE 397 AA; 43515 MW; AACEE63869ED8F3 CRC64;

Query Match 96.6%; Score 28; DB 10; Length 397;
 Best Local Similarity 83.3%; Pred. No. 2.1e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GHXXKL 6
 |||||
 Db 283 GHKTKL 288

RESULT 9

Q8N7X4 ID Q8N7X4 PRELIMINARY; PRT; 407 AA.
 AC Q8N7X4
 DT 01-OCT-2002 (TREMBLrel. 22, Created)
 DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
 DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
 DE Hypothetical protein FLJ40242.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kawakami B., Sugiyama A., Takemoto M., Sugiyama T., Irie R.,
 RA Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,
 RA Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,
 RA Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K., Nagatsuma M.,
 RA Murakawa K., Kanehori K., Takahashi-Fujii A., Oshima A., Suzuki Y.,
 RA Sugano S., Nagahari K., Masuho Y., Nagai K., Isogai T.;
 RT "NEDO human cDNA sequencing project."
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AK097561; BAC05102.1;
 DR InterPro; IPR002190; MAGE.

DR Pfam; PF01454; MAGE; 1.
 DR PROSITE; PS00838; MAGE; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 407 AA; 44091 MW; D674F426131244C0 CRC64;

Query Match 96.6%; Score 28; DB 4; Length 407;
 Best Local Similarity 83.3%; Pred. No. 2.1e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GHXXKL 6
 |||||
 Db 4 GHKSKL 9

RESULT 10

Q9PXU3 ID Q9PXU3 PRELIMINARY; PRT; 522 AA.
 AC Q9PXU3
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
 DE GAG polyprotein [Contains: core protein(s) P24].
 GN GAG.
 OS Chimpanzee immunodeficiency virus (SIV(cpz)) (CIV).
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11723;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94025594; PubMed=8212578;
 RA Hirsch V.M., McGann C., Dapolito G., Goldstein S., Ogen-Odoi A.,
 RA Biriyawho B., Lakwo T., Johnson P.R.;
 RT "Identification of a new subgroup of SIVagm in tantalus monkeys.";
 RL Virology 197:426-430(1993).
 DR HSP; P03351; 2E1A.
 DR GO; GO:0019012; C:virion; IEA.
 DR GO; GO:0003676; F:nucleic acid binding; IEA.
 DR GO; GO:0005198; F:structural molecule activity; IEA.
 DR GO; GO:0016032; P:viral life cycle; IEA.
 DR InterPro; IPR000721; Gag_P24.
 DR InterPro; IPR000071; Retrovir_p17.
 DR InterPro; IPR008916; Retrov_capsid_C.
 DR InterPro; IPR008919; Retrov_capsid_N.
 DR InterPro; IPR001878; Znf_CCHC.
 DR Pfam; PF00540; Gag_p17; 1.
 DR Pfam; PF00607; Gag_P24; 1.
 DR Pfam; PF00098; zf-CCHC; 2.
 DR PRINTS; PR00939; C2HCZNFINGER.
 DR PRINTS; PR00234; HIVMATRIX.
 DR SMART; SM00343; ZNF_C2HC; 2.
 DR PROSITE; PS0158; ZF_CCHC; 2.
 KW AIDS; Core protein; Polyprotein.
 SQ SEQUENCE 522 AA; 57956 MW; 81E5B654D69D1F CRC64;

Query Match 96.6%; Score 28; DB 15; Length 522;
 Best Local Similarity 83.3%; Pred. No. 2.7e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GHXXKL 6
 |||||
 Db 366 GHKAKL 371

RESULT 11

P89903 ID P89903 PRELIMINARY; PRT; 523 AA.
 AC P89903
 DT 01-MAY-1997 (TREMBLrel. 03, Created)
 DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
 DE GAG polyprotein.
 GN GAG.
 OS Chimpanzee immunodeficiency virus (SIV(cpz)) (CIV).
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.

```

OX NCBI_TaxID=11723;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=tantalus-1;
RX MEDLINE=97223354; PubMed=9123848;
RA Soares M.A., Robertson D.L., Hui H., Allan J.S., Shaw G.M., Hahn B.H.;
RT "A full-length and replication-competent proviral clone of SIVAGM from
  tantalus monkeys.";
RL Virology 228:394-399 (1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=tantalus-1;
RX STRAIN=tantalus-1;
RA Soares M.;
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U5991; AAC57051.1; -.
DR HSSP; P03351; IEIA.
DR GO; GO:0019012; C:virion; IEA.
DR GO; GO:0003676; P:nucleic acid binding; IEA.
DR GO; GO:0005198; P:structural molecule activity; IEA.
DR GO; GO:0016032; P:viral life cycle; IEA.
DR InterPro; IPR000721; Gag_P24.
DR InterPro; IPR000071; Retrovir_P17.
DR InterPro; IPR008916; Retrov_capsid_C.
DR InterPro; IPR008919; Retrov_capsid_N.
DR InterPro; IPR001878; Znf_CCHC.
DR Pfam; PF00540; Gag_P17; 1.
DR Pfam; PF00607; Gag_P24; 1.
DR Pfam; PF00098; zf_CCHC; 2.
DR PRINTS; PR00939; C2HCZNFINGER.
DR PRINTS; PR00234; HIV1NATRIX.
DR SMART; SM00343; ZNF_C2HC; 2.
DR PROSITE; PS0158; ZF_CCHC; 2.
KW AIDS; Core protein; Polyprotein.
SQ SEQUENCE 523 AA; 57688 MW; F623CED7277681AA CRC64;

Query Match 96.6%; Score 28; DB 15; Length 523;
Best Local Similarity 83.3%; Pred. No. 2.8e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GHKXKL 6
DB 366 GHKAKL 371

RESULT 12
P91451
ID P91451 PRELIMINARY; PRT; 786 AA.
AC P91451;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN T16A1.2.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peleodermidae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RL "Genome sequence of the nematode C. elegans: a platform for
  investigating biology. The C. elegans Sequencing Consortium.";
  Science 282:2012-2018 (1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX Du Z., Goela D.;
RA "The sequence of C. elegans cosmid T16A1.";
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.

```

```

RC STRAIN=Bristol N2;
RA Waterston R.;
RT "Direct Submission.";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; U80454; AAB37879.1; -.
DR PIR; G88065; G88065.
DR WormPep; T16A1.2; CE13706.
DR InterPro; IPR003125; WSN.
DR Pfam; PF02206; WSN; 1.
DR SMART; SM00453; WSN; 1.
KW Hypothetical protein.
SQ SEQUENCE 786 AA; 90235 MW; 831AAB0269B2E9A7 CRC64;

Query Match 96.6%; Score 28; DB 5; Length 786;
Best Local Similarity 83.3%; Pred. No. 4.1e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GHKXKL 6
DB 17 GHKTKL 22

RESULT 13
Q9PT70
ID Q9PT70 PRELIMINARY; PRT; 867 AA.
AC Q9PT70;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE FCPI serine phosphatase.
GN FCPI.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RA Golini E., Marazziti D., Matteoni R., Tocchini-Valentini G.P.;
RT "Molecular cloning of the Xenopus laevis FCPI serine phosphatase
  cDNA.";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ132385; CAB65510.1; -.
DR GO; GO:0005622; C:intracellular; IEA.
DR InterPro; IPR001357; BRCT.
DR InterPro; IPR004274; NIF.
DR Pfam; PF00533; BRCT; 1.
DR Pfam; PF03031; NIF; 1.
DR SMART; SM00292; BRCT; 1.
DR SMART; SM00577; CPDC; 1.
DR PROSITE; PS0172; BRCT; 1.
SQ SEQUENCE 867 AA; 97438 MW; B5126AB886A4F2E1 CRC64;

Query Match 96.6%; Score 28; DB 13; Length 867;
Best Local Similarity 83.3%; Pred. No. 4.6e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GHKXKL 6
DB 820 GHKAKL 825

RESULT 14
Q9KM19
ID Q9KM19 PRELIMINARY; PRT; 103 AA.
AC Q9KM19;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Sulf family protein.
GN VCA0570.
OS Vibrio cholerae.

```

```

OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=El Tor N16961 / Serotype O1;
RX MEDLINE=20406833; PubMed=10952301;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Emolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT cholerae.";
RL Nature 406:477-483(2000).
DR EMBL; AE004388; AAF96472.1; -.
DR PIR; G82442; G82442.
DR HSP; P08245; 1D1R.
DR TIGR; VCA0570; -.
DR GO; GO:0003743; F:translation initiation factor activity; IEA.
DR GO; GO:0006413; P:translational initiation; IEA.
DR InterPro; IPR003872; SUI1_arch_bact.
DR InterPro; IPR001950; TIF_SUI1.
DR Pfam; PF01253; SUI1; 1.
DR TIGRFAMs; TIGR01158; SUI1_rel; 1.
DR PROSITE; PS01118; SUI1_1; 1.
DR PROSITE; PS02096; SUI1_2; 1.
KW Complete proteome.
SQ SEQUENCE 103 AA; 11124 MW; 57388B4E8E77762 CRC64;

Query Match 93.1%; Score 27; DB 16; Length 103;
Best Local Similarity 83.3%; Pred. No. 93;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GHXXKL 6
Db 95 GHVKVL 100

RESULT 15
Q82U99 Q82U99 PRELIMINARY; PRT; 123 AA.
AC Q82U99;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein.
GN NE1597.
OS Nitrosomonas europaea.
OC Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales;
OC Nitrosomonadaceae; Nitrosomonas.
OX NCBI_TaxID=915;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 19718 /IFO 14298;
RX MEDLINE=22586410; PubMed=12700255;
RA Chain P., Lamerdin J.E., Larimer F.W., Regala W., Lao V., Land M.,
RA Hauser L., Hooper A.B., Klotz M.G., Norton J., Sayavedra-Soto L.A.,
RA Arciero D.M., Hommes N.G., Whittaker M.M., Arp D.J.;
RT "Complete genome sequence of the ammonia-oxidizing bacterium and
RL J. Bacteriol. 185:2759-2773(2003).
DR EMBL; BX321861; CAD95508.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 123 AA; 14233 MW; E8A400B7FD43C05 CRC64;

Query Match 93.1%; Score 27; DB 16; Length 123;
Best Local Similarity 83.3%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GHXXKL 6

```

Db 110 GHKKL 115

Search completed: March 4, 2004, 13:10:31
Job time : 13.2941 secs

GenCore version 5.1.6

Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2004, 13:04:08 ; Search time 13 Seconds
(without alignments)
67.511 Million cell updates/sec

Title: SEQ-A

Perfect score: 98

Sequence: 1 HGHEQQHGLGHGKXKL 17

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:*
1: /cgn2_6/ptodata/2/iaa/5A.COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B.COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A.COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B.COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/6C.COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/BACKFILES.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	96	98.0	55	3	US-09-612-126-5
2	96	98.0	62	3	US-09-612-126-7
3	96	98.0	83	3	US-09-612-126-6
4	96	98.0	94	3	US-09-612-126-10
5	96	98.0	186	3	US-09-612-126-8
6	96	98.0	255	3	US-09-612-126-1
7	74	75.5	179	3	US-09-612-126-11
8	69	70.4	15	3	US-09-612-126-2
9	69	70.4	16	3	US-09-612-126-9
10	61.5	62.8	47	3	US-09-612-126-4
11	59	60.2	398	3	US-09-461-474-17
12	57	58.2	474	3	US-09-461-474-10
13	53	54.1	448	3	US-09-461-474-8
14	51	52.0	21	3	US-09-612-126-3
15	49	50.0	11	3	US-09-612-126-12
16	49	50.0	77	4	US-09-252-991A-19317
17	49	50.0	788	4	US-09-252-991A-17380
18	48.5	49.5	572	4	US-09-252-991A-17898
19	48.5	49.5	680	4	US-09-252-991A-29223
20	47	48.0	249	4	US-09-543-681A-6793
21	47	48.0	363	4	US-09-528-352-4930
22	47	48.0	400	4	US-09-543-681A-6151
23	47	48.0	546	4	US-09-252-991A-30078
24	46	46.9	191	4	US-09-252-991A-29683
25	46	46.9	264	4	US-09-134-000C-4103
26	46	46.9	344	4	US-09-134-001C-3524
27	46	46.9	500	4	US-09-252-991A-31980

```

28 45.5 46.4 218 4 US-09-252-991A-25291 Sequence 25291, A
29 45 342 4 US-09-134-001C-4190 Sequence 4190, Ap
30 45 538 4 US-09-252-991A-21622 Sequence 21622, A
31 44.5 45.4 293 4 US-09-461-325-233 Sequence 233, App
32 44.5 45.4 293 4 US-10-012-542-233 Sequence 233, App
33 44.5 45.4 307 4 US-09-461-325-331 Sequence 331, App
34 44.5 45.4 307 4 US-10-012-542-331 Sequence 331, App
35 44.5 45.4 490 4 US-09-461-325-158 Sequence 158, App
36 44.5 45.4 490 4 US-10-012-542-158 Sequence 158, App
37 44.5 45.4 504 4 US-09-461-325-332 Sequence 332, App
38 44.5 45.4 504 4 US-10-012-542-332 Sequence 332, App
39 44.5 45.4 607 4 US-09-252-991A-20596 Sequence 20596, A
40 44 226 4 US-09-543-681A-5920 Sequence 5920, Ap
41 44 256 4 US-09-252-991A-19839 Sequence 19839, A
42 44 274 4 US-09-711-164-369 Sequence 369, App
43 44 274 4 US-09-711-164-407 Sequence 407, App
44 44 404 4 US-09-252-991A-18670 Sequence 18670, A
45 43.5 44.4 60 1 US-08-255-457-1 Sequence 1, Appli

```

ALIGNMENTS

```

RESULT 1
US-09-612-126-5
; Sequence 5, Application US/09612126
; Patent No. 6284726
; GENERAL INFORMATION:
; APPLICANT: TEMPLE UNIVERSITY - OF THE COMMONWEALTH SYSTEM OF
; TITLE OF INVENTION: INHIBITION OF ANGIOGENESIS BY PEPTIDE ANALOGS OF HIGH
; TITLE OF INVENTION: MOLECULAR WEIGHT KININOGEN DOMAN 5
; FILE REFERENCE: 6056-258 CTI
; CURRENT APPLICATION NUMBER: US/09/612,126
; CURRENT FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/187,844
; PRIOR FILING DATE: 1998-11-10
; PRIOR APPLICATION NUMBER: PCT/US99/26377
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 55
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Human high
; OTHER INFORMATION: molecular weight kininogen light chain amino acids
; OTHER INFORMATION: Lys (420) through Asp (474)
US-09-612-126-5

```

```

Query Match 98.0%; Score 96; DB 3; Length 55;
Best Local Similarity 94.1%; Pred. No. 1.9e-07;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 HGHEQQHGLGHGKXKL 17
Db 26 HGHEQQHGLGHGKFKL 42

```

```

RESULT 2
US-09-612-126-7
; Sequence 7, Application US/09612126
; Patent No. 6284726
; GENERAL INFORMATION:
; APPLICANT: TEMPLE UNIVERSITY - OF THE COMMONWEALTH SYSTEM OF
; TITLE OF INVENTION: INHIBITION OF ANGIOGENESIS BY PEPTIDE ANALOGS OF HIGH
; TITLE OF INVENTION: MOLECULAR WEIGHT KININOGEN DOMAN 5
; FILE REFERENCE: 6056-258 CTI
; CURRENT APPLICATION NUMBER: US/09/612,126
; CURRENT FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/107,844
; PRIOR FILING DATE: 1998-11-10
; PRIOR APPLICATION NUMBER: PCT/US99/26377

```



```
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 62
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Human high
; OTHER INFORMATION: molecular weight kininogen light chain amino acids
; OTHER INFORMATION: His(441) through Lys(502)
US-09-612-126-7

Query Match          98.0%; Score 96; DB 3; Length 62;
Best Local Similarity 94.1%; Pred. No. 2.1e-07;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 HGHEQQHGLGHGKXKL 17
   |||||
Db 5 HGHEQQHGLGHGKFKL 21

RESULT 3
US-09-612-126-6
; Sequence 6, Application US/09612126
; Patent No. 6284726
; GENERAL INFORMATION:
; APPLICANT: TEMPLE UNIVERSITY - OF THE COMMONWEALTH SYSTEM OF
; TITLE OF INVENTION: INHIBITION OF ANGIOGENESIS BY PEPTIDE ANALOGS OF HIGH
; TITLE OF INVENTION: MOLECULAR WEIGHT KININOGEN DOMAN 5
; FILE REFERENCE: 6056-258 CTI
; CURRENT APPLICATION NUMBER: US/09/612,126
; CURRENT FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/107,844
; PRIOR FILING DATE: 1998-11-10
; PRIOR APPLICATION NUMBER: PCT/US99/26377
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 83
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Human high
; OTHER INFORMATION: molecular weight kininogen light chain amino acids
; OTHER INFORMATION: Lys(420) through Lys(502)
US-09-612-126-6

Query Match          98.0%; Score 96; DB 3; Length 83;
Best Local Similarity 94.1%; Pred. No. 2.9e-07;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 HGHEQQHGLGHGKXKL 17
   |||||
Db 26 HGHEQQHGLGHGKFKL 42

RESULT 4
US-09-612-126-10
; Sequence 10, Application US/09612126
; Patent No. 6284726
; GENERAL INFORMATION:
; APPLICANT: TEMPLE UNIVERSITY - OF THE COMMONWEALTH SYSTEM OF
; TITLE OF INVENTION: INHIBITION OF ANGIOGENESIS BY PEPTIDE ANALOGS OF HIGH
; TITLE OF INVENTION: MOLECULAR WEIGHT KININOGEN DOMAN 5
; FILE REFERENCE: 6056-258 CTI
; CURRENT APPLICATION NUMBER: US/09/612,126
; CURRENT FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/107,844
; PRIOR FILING DATE: 1998-11-10
; PRIOR APPLICATION NUMBER: PCT/US99/26377
; PRIOR FILING DATE: 1999-11-09
```

```
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 94
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Human high
; OTHER INFORMATION: molecular weight kininogen light chain amino acids
; OTHER INFORMATION: Lys(420) through Ser(513)
US-09-612-126-10

Query Match          98.0%; Score 96; DB 3; Length 94;
Best Local Similarity 94.1%; Pred. No. 3.2e-07;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 HGHEQQHGLGHGKXKL 17
   |||||
Db 26 HGHEQQHGLGHGKFKL 42

RESULT 5
US-09-612-126-8
; Sequence 8, Application US/09612126
; Patent No. 6284726
; GENERAL INFORMATION:
; APPLICANT: TEMPLE UNIVERSITY - OF THE COMMONWEALTH SYSTEM OF
; TITLE OF INVENTION: INHIBITION OF ANGIOGENESIS BY PEPTIDE ANALOGS OF HIGH
; TITLE OF INVENTION: MOLECULAR WEIGHT KININOGEN DOMAN 5
; FILE REFERENCE: 6056-258 CTI
; CURRENT APPLICATION NUMBER: US/09/612,126
; CURRENT FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/107,844
; PRIOR FILING DATE: 1998-11-10
; PRIOR APPLICATION NUMBER: PCT/US99/26377
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 186
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Human high
; OTHER INFORMATION: molecular weight kininogen light chain amino acids
; OTHER INFORMATION: His(441) through Ser(626)
US-09-612-126-8

Query Match          98.0%; Score 96; DB 3; Length 186;
Best Local Similarity 94.1%; Pred. No. 6.4e-07;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 HGHEQQHGLGHGKXKL 17
   |||||
Db 5 HGHEQQHGLGHGKFKL 21

RESULT 6
US-09-612-126-1
; Sequence 1, Application US/09612126
; Patent No. 6284726
; GENERAL INFORMATION:
; APPLICANT: TEMPLE UNIVERSITY - OF THE COMMONWEALTH SYSTEM OF
; TITLE OF INVENTION: INHIBITION OF ANGIOGENESIS BY PEPTIDE ANALOGS OF HIGH
; TITLE OF INVENTION: MOLECULAR WEIGHT KININOGEN DOMAN 5
; FILE REFERENCE: 6056-258 CTI
; CURRENT APPLICATION NUMBER: US/09/612,126
; CURRENT FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/107,844
; PRIOR FILING DATE: 1998-11-10
; PRIOR APPLICATION NUMBER: PCT/US99/26377
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 12
```

```

RESULT 10
US-09-612-126-4
; Sequence 4, Application US/09612126
; Patent No. 6284726
; GENERAL INFORMATION:
; APPLICANT: TEMPLE UNIVERSITY - OF THE COMMONWEALTH SYSTEM OF
; TITLE OF INVENTION: INHIBITION OF ANGIOGENESIS BY PEPTIDE ANALOGS OF HIGH
; FILE REFERENCE: 6056-258 CTI
; CURRENT APPLICATION NUMBER: US/09/612,126
; CURRENT FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/107,844
; PRIOR FILING DATE: 1998-11-10
; PRIOR APPLICATION NUMBER: PCT/US99/26377
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 47
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:

```

; OTHER INFORMATION: Description of Artificial Sequence: Human high
 ; OTHER INFORMATION: molecular weight kininogen light chain amino acids
 ; OTHER INFORMATION: Gly(456) through Lys(502)
 US-09-612-126-4

Query Match 62.8%; Score 61.5; DB 3; Length 47;
 Best Local Similarity 64.7%; Pred. No. 0.016;
 Matches 11; Conservative 2; Mismatches 3; Indels 1; Gaps 1;

QY 1 HGHEQQHGLGHG-HKXK 16
 ||||: ||||| |||||
 DB 20 HGKHKEGHGKHGKHKK 36

RESULT 11

US-09-461-474-17
 ; Sequence 17, Application US/09461474
 ; Patent No. 6278042
 ; GENERAL INFORMATION:
 ; APPLICANT: Allen, Steve
 ; APPLICANT: Rafalski, Antoni
 ; APPLICANT: Sakai, Hajime
 ; TITLE OF INVENTION: Plant Metal Transporters
 ; FILE REFERENCE: BB1303 US NA
 ; CURRENT APPLICATION NUMBER: US/09/461.474
 ; CURRENT FILING DATE: 1999-12-14
 ; EARLIER APPLICATION NUMBER: 60/112.562
 ; EARLIER FILING DATE: 1998-12-16
 ; NUMBER OF SEQ ID NOS: 17
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO 17
 ; LENGTH: 398
 ; TYPE: PRT
 ; ORGANISM: Arabidopsis thaliana
 US-09-461-474-17

Query Match 60.2%; Score 59; DB 3; Length 398;
 Best Local Similarity 69.2%; Pred. No. 0.32;
 Matches 9; Conservative 0; Mismatches 4; Indels 4; Gaps 0;

QY 1 HGHEQQHGLGHG 13
 ||||: ||||| |||||
 DB 184 HGKHKEGHGKHGKH 196

RESULT 12

US-09-461-474-10
 ; Sequence 10, Application US/09461474
 ; Patent No. 6278042
 ; GENERAL INFORMATION:
 ; APPLICANT: Allen, Steve
 ; APPLICANT: Rafalski, Antoni
 ; APPLICANT: Sakai, Hajime
 ; TITLE OF INVENTION: Plant Metal Transporters
 ; FILE REFERENCE: BB1303 US NA
 ; CURRENT APPLICATION NUMBER: US/09/461.474
 ; CURRENT FILING DATE: 1999-12-14
 ; EARLIER APPLICATION NUMBER: 60/112.562
 ; EARLIER FILING DATE: 1998-12-16
 ; NUMBER OF SEQ ID NOS: 17
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO 10
 ; LENGTH: 474
 ; TYPE: PRT
 ; ORGANISM: Oryza sativa
 US-09-461-474-10

Query Match 58.2%; Score 57; DB 3; Length 474;
 Best Local Similarity 69.2%; Pred. No. 0.75;
 Matches 9; Conservative 0; Mismatches 4; Indels 4; Gaps 0;

QY 1 HGHEQQHGLGHG 13
 ||||: ||||| |||||

DB 240 HGKHKEGHGKHGKH 252

RESULT 13

US-09-461-474-8
 ; Sequence 8, Application US/09461474
 ; Patent No. 6278042
 ; GENERAL INFORMATION:
 ; APPLICANT: Allen, Steve
 ; APPLICANT: Rafalski, Antoni
 ; APPLICANT: Sakai, Hajime
 ; TITLE OF INVENTION: Plant Metal Transporters
 ; FILE REFERENCE: BB1303 US NA
 ; CURRENT APPLICATION NUMBER: US/09/461.474
 ; CURRENT FILING DATE: 1999-12-14
 ; EARLIER APPLICATION NUMBER: 60/112.562
 ; EARLIER FILING DATE: 1998-12-16
 ; NUMBER OF SEQ ID NOS: 17
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO 8
 ; LENGTH: 448
 ; TYPE: PRT
 ; ORGANISM: Zea mays
 US-09-461-474-8

Query Match 54.1%; Score 53; DB 3; Length 448;
 Best Local Similarity 66.7%; Pred. No. 2.7;
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 HGHEQQHGLGHG 13
 ||||: ||||| |||||

DB 215 GHGKHKEGHGKHGKH 226

RESULT 14

US-09-612-126-3
 ; Sequence 3, Application US/09612126
 ; Patent No. 6284726
 ; GENERAL INFORMATION:
 ; APPLICANT: TEMPLE UNIVERSITY - OF THE COMMONWEALTH SYSTEM OF
 ; TITLE OF INVENTION: INHIBITION OF ANGIOGENESIS BY PEPTIDE ANALOGS OF HIGH
 ; TITLE OF INVENTION: MOLECULAR WEIGHT KININOGEN DOMAN 5
 ; FILE REFERENCE: 6056-258 CTI
 ; CURRENT APPLICATION NUMBER: US/09/612.126
 ; CURRENT FILING DATE: 2000-07-07
 ; PRIOR APPLICATION NUMBER: 60/107,844
 ; PRIOR FILING DATE: 1998-11-10
 ; PRIOR APPLICATION NUMBER: PCT/US99/26377
 ; PRIOR FILING DATE: 1999-11-09
 ; NUMBER OF SEQ ID NOS: 12
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 3
 ; LENGTH: 21
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: Human high
 ; OTHER INFORMATION: molecular weight kininogen light chain amino acids
 ; OTHER INFORMATION: Lys(420) through Gly(440)
 US-09-612-126-3

Query Match 52.0%; Score 51; DB 3; Length 21;
 Best Local Similarity 50.0%; Pred. No. 0.24;
 Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 HGHEQQHGLGHGKH 14
 ||||: ||||| |||||

DB 6 HGKHKEGHGKHGKH 19

RESULT 15

US-09-612-126-12
 ; Sequence 12, Application US/09612126

```

; Patent No. 6284726
; GENERAL INFORMATION:
; APPLICANT: TEMPLE UNIVERSITY - OF THE COMMONWEALTH SYSTEM OF
; TITLE OF INVENTION: INHIBITION OF ANGIOGENESIS BY PEPTIDE ANALOGS OF HIGH
; FILE REFERENCE: 6056-258 CT1
; CURRENT APPLICATION NUMBER: US/09/612,126
; CURRENT FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/107,844
; PRIOR FILING DATE: 1998-11-10
; PRIOR APPLICATION NUMBER: PCT/US99/26377
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 12
; LENGTH: 11
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Human high
; OTHER INFORMATION: molecular weight kininogen light chain amino acids
; OTHER INFORMATION: His(475) through His(485)
US-09-612-126-12

```

```

Query Match      50.0%; Score 49; DB 3; Length 11;
Best Local Similarity 63.6%; Pred. No. 0.25;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 1 HGHEQOHGLGH 11
   |||.:|||
Db 1 HGHKKKHGH 11

```

```

Search completed: March 4, 2004, 13:12:26
Job time : 14 secs

```